VOL. XIII APRIL, 1929 No. 2

# *The Journal of* APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University, Athens, Ohio
WITH A BOARD OF CO-OPERATING EDITORS

## TABLE OF CONTENTS

H. H. REMMERS. The Measurement of Interest Differences between Students of Engineering and of Agriculture ..... 105
MILES A. TINKER AND DONALD G. PATERSON. Studies of Typographical Factors Influencing Speed of Reading. II. Size of Type ..... 120
S. D. PORTEUS. The Social Psychology of the Australian Aboriginal ..... 131
LOUISE E. POULL AND RUTH P. MONTGOMERY. The Porteus Maze Test as a Discriminative Measure in Delinquency ..... 145
E. J. ASHER. The Predictive Value of Mental Tests that Satisfy Spearman's Tetrad Criterion ..... 152
M. EUSTACE BROOM, BLANCHE THOMPSON AND M. THELMA BOUTON. Sex Difference in Handwriting ..... 159
KARL C. GARRISON. An Investigation of Some Simple Speed Activities ..... 167
HOWARD P. LONGSTAFF AND JAMES P. PORTER. Objective Measures of Ability and Achievement in General Psychology ..... 173
NORMA V. SCHEIDEMANN. An Experiment in Teaching Psychology ..... 188
NEWS AND NOTES ..... 192
BOOK REVIEWS ..... 195
NEW BOOKS AND PAMPHLETS ..... 204

PUBLISHED BI-MONTELY BY THE EDITOR
All communications should be addressed to the Editor,
Ohio University, Athens, Ohio
Printed by
WAVERLY PRESS, INC.
Baltimore, Md.

Entered as second class matter June 3, 1924, at the post office at Baltimore, Maryland, under the act of March 3, 1879.
Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

COPYRIGHT 1929, BY JAMES P. PORTER

# The Journal of Applied Psychology


## CO-OPERATING EDITORS

W. F. Book, Indiana University; W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Bureau of Special Education, 939 Harrison Avenue, Columbus, Ohio; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.


The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.


The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

*Errata—February Issue*

A series of mistakes led to the omission of the name of the reviewer of the books by Professors Murchison and Root in the February issue. This reviewing was done by the person whose name appears below. It is requested that all subscribers use the part of this slip carrying the name to paste in where the name of the reviewer should appear in the February issue of the Journal (page 102).

SARAH C. FISHER,
*University of California at Los Angeles*

# THE MEASUREMENT OF INTEREST DIFFERENCES BETWEEN STUDENTS OF ENGINEERING AND OF AGRICULTURE

H. H. REMMERS

*Purdue University*

## I. INTRODUCTORY

Within recent years there has been published a growing volume of studies concerned with the measurement of vocational-educational interests. Thorndike (23, 24) in 1912 reported a study of the relation of interest to ability in various school subjects for a group of Barnard College students. He reported (*a*) a considerable permanence of pupil's interests; (*b*) an equal permanence of abilities; and (*c*) a high correlation (0.89) between interest and ability.

Bridges and Dollinger (2) in a study designed to check Thorndike's conclusion concerning the close relationship of interest and ability conclude that "a person's relative interests are an extraordinarily inaccurate symptom of his relative capacities." In an evaluation of this study Thorndike (25) concludes that on the whole when unreliability of school marks is allowed for the data of Bridges and Dollinger seem rather to corroborate his own findings. King and Adelstein (13) using Thorndike's questionnaire method with 140 University of Iowa students obtained median correlations somewhat lower than those reported by Thorndike.

Fryer (8) in 1923–24 concludes on the basis of data collected at the University of Utah and New York University that while there is a substantial relationship between educational interests and estimated abilities ($r = 0.70$), there is very little relationship when school marks instead of estimates of ability are used.

A. R. Crathorne (4) investigated the vocational interests of

2000 college freshmen in a number of institutions again by the questionnaire method. He found that on entering high school 57 per cent had decided upon their life work but that by the time they reached college almost exactly 50 per cent of these had altered their choices. Eighty-seven per cent of college freshmen had arrived at definite vocational interests by the time they entered college.

The highly subjective nature of such studies as the foregoing leaves conclusions based upon them rather questionable. A series of studies by Moore (16, 17), Freyd (5, 6), Cowdery (3), Miner (14, 15), Ream (19), Strong (20, 21, 22), Hubbard (16) and others introduced new techniques and methods of greater refinement both in the measurement of interests and statistical treatment of the data. Without attempting to summarize the findings of the various studies, it is sufficient to say that more reliable and useful results were obtained—useful, that is, for those interested in educational-vocational counseling, guidance, and selection. The study to be reported here grew out of the problem of such possible guidance at Purdue University.

## II. THE PROBLEM

Is it possible to construct a measuring instrument which will discriminate between students with interests predominantly in engineering and those with interests primarily in agriculture? The problem is obviously one of considerable practical value for applied psychology.

Every year there is an appreciable number of freshmen who come to our universities with various sorts of misconceptions about the sort of life work which they wish to follow. The "radio halo" in electrical engineering is a case in point. A boy whose aptitudes, temperament, and interests fit him admirably for an agricultural career builds a radio set according to directions (cook-book style) and is convinced as a result that he wishes to become an electrical engineer. When, however, he comes to the technical subjects which are necessary prerequisite for such a vocation, he frequently finds that his enthusiasm was unwarranted in the light of his experience with courses in mathematics, engineering problems, and the like.

Were there available at the beginning of the student's university career even a fairly reliable index of his vocational interests supplementary to indices of mental capacity, scholastic aptitudes, and such other measures as may be related to vocational success or failure, it would be a valuable additional tool in the armamentarium of the teacher or the vocational counseler. It is of course obvious that an accurate notion of a student's *real* vocational interests will be a very significant factor in his relative success or failure, and it is important, therefore, to know quickly and accurately when a boy is interviewed just where these interests lie. It was such considerations as these which led to the study here reported.

### III. PROCEDURE

During the orientation period for freshmen at Purdue University in the fall of 1926 there was administered the Purdue Interest Report Blank devised by the writer. The data thus obtained constitute the basis of the present study. A sample of the Blank, which required 50 minutes to administer, is given below.

| Very strong dislike | Marked dislike | Some dislike | Indifferent | Some liking | Marked liking | Very strong liking |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Above is a scale of values in terms of which you are to judge the items listed below. If you dislike an item very strongly, mark it 1; if you like it very strongly, mark it 7; if you neither like or dislike it, mark it 4, and so for any of the intermediate values.

| | |
|---|---|
| Actor..... — | Aviator..... — |
| Architect..... — | Bank Teller..... — |
| Army Officer..... — | Locomotive Engineer.... — |
| Artist..... — | Machinist..... — |
| Astronomer..... — | Magazine Writer..... — |
| Athletic Director..... — | Marine Engineer..... — |
| Auctioneer..... — | Mechanical Engineer..... — |
| Auto Salesman..... — | Missionary..... — |
| Auto Repairman..... — | Musician..... — |

| | | | |
|---|---|---|---|
| Newspaper Reporter | — | Brunettes | — |
| Naval Officer | — | Very old people | — |
| Novelist | — | Children | — |
| Office Clerk | — | Cripples | — |
| Office Manager | — | Side-show freaks | — |
| Short women | — | People with gold teeth | — |
| Blondes | — | People with receding teeth | — |

The 349 items for this instrument were obtained from various sources. (1) An Interest Blank in use at Stanford University was kindly furnished by Dr. Karl M. Cowdery. The entire list of items in this blank was included. (2) Paired antithetical statements of interests were furnished by Professor R. B. Miner of the University of Kentucky. A number of these which did not overlap the Stanford items were adapted and included. (3) Forty-five items presumably related to agricultural interests were furnished by Professors M. L. Fisher and E. C. Young of the Purdue School of Agriculture.

It was thought at first that degrees of like or dislike relative to particular items might be of significance. Inspection of the item tabulation of 100 blanks selected at random for students of agriculture and of 200 blanks for students of engineering showed, however, that this hypothesis was not borne out. Items marked with any number less than 4 were, therefore, fused under the general caption of "dislike" and those above 4 under "like." Those marked 4 were disregarded. The percentage of "like" and "dislike" together with the probable error[1] was computed from the formula

$$\text{P.E.}_p = 0.6745\sqrt{\frac{pq}{n}}$$

[1] See Yule, G. U. An Introduction to the Theory of Statistics, p. 257, formula (2), 1924. Fortunately there was available the "Table of Standard Errors and Probable Errors of Percentages for Varying Numbers of Cases" by H. A. Edgerton and Donald G. Paterson (see Jour. Applied Psychol., x, 3, September, 1926, pp. 378–391). Without the aid of such a table the statistical labor involved in the present study would have been well nigh prohibitive.

TABLE 1

*Differential items for students of agriculture and engineering, on Purdue Interest Questionnaire*

| ITEM | | Per cent Diff. Like | P.E. diff. | Diff. / P.E. diff. | Per cent Diff. Dislike | P.E. diff. | Diff. / P.E. diff. |
|---|---|---|---|---|---|---|---|
| 1 | Actor | 14 E | 3.4 | 4.1 | 8 A | 3.9 | 2.0 |
| 2 | Architect | 35 E | 3.8 | 9.2 | 17 A | 3.3 | 5.2 |
| 4 | Artist | | | | 11 A | 4.1 | 2.7 |
| 6 | Athletic Director | 20 A | 3.9 | 5.1 | | | |
| 7 | Auctioneer | | | | 9 E | 4.0 | 2.5 |
| 9 | Auto Racer | 12 E | 3.9 | 3.1 | 10 A | 4.0 | 2.5 |
| 11 | Aviator | 22 E | 4.1 | 5.4 | 14 A | 3.3 | 4.2 |
| 13 | Baseball player | | | | 10 E | 3.5 | 2.9 |
| 15 | Building Contractor | 31 E | 3.9 | 7.9 | | | |
| 21 | Chemist | 15 E | 4.1 | 3.7 | 13 A | 3.8 | 3.4 |
| 26 | Draftsman | 41 E | 3.5 | 11.7 | 14 A | 3.4 | 4.1 |
| 27 | Editor | | | | 10 A | 3.9 | 2.6 |
| 28 | Explorer | 9 E | 4.1 | 2.2 | 10 A | 3.2 | 3.1 |
| 29 | Factory Manager | 23 E | 4.0 | 5.8 | 11 A | 3.5 | 3.1 |
| 31 | Farmer | 63 A | 3.1 | 20.0 | 47 E | 3.1 | 15.2 |
| 33 | Floorwalker | | | | 8 E | 3.9 | 2.1 |
| 36 | Gardener | 35 A | 3.7 | 9.7 | 26 E | 3.8 | 6.8 |
| 39 | Judge | 8 E | 3.3 | 2.4 | 11 A | 4.1 | 2.7 |
| 40 | Labor Arbitrator | | | | 10 A | 4.1 | 2.4 |
| 41 | Landscape Gardener | 23 A | 3.9 | 5.9 | 16 E | 3.8 | 4.2 |
| 45 | Locomotive Engineer | 16 E | 3.9 | 4.1 | 12 A | 3.6 | 3.3 |
| 46 | Machinist | 19 E | 3.7 | 5.1 | | | |
| 47 | Magazine | | | | 9 A | 4.1 | 2.2 |
| 48 | Marine Engineer | 39 E | 3.9 | 10.0 | 22 A | 3.5 | 6.3 |
| 49 | Mechanical Engineer | | | | 20 A | 3.1 | 6.4 |
| 53 | Naval Officer | 11 E | 3.5 | 3.1 | 11 A | 3.8 | 2.9 |
| 58 | Pharmacist | 15 E | 4.1 | 3.7 | 15 A | 4.0 | 3.8 |
| 66 | Rancher | 22 A | 3.6 | 6.1 | | | |
| 68 | Scientific Research Worker | 21 E | 3.9 | 5.4 | 17 A | 3.6 | 4.9 |
| 71 | Shop Foreman | 21 E | 3.9 | 5.4 | 14 A | 3.7 | 3.8 |
| 72 | Social Worker | | | | 14 E | 4.1 | 3.4 |
| 76 | Stock Broker | 9 E | 3.8 | 2.4 | 22 A | 4.1 | 5.4 |
| 79 | Surgeon | 12 E | 3.7 | 3.2 | 13 A | 4.1 | 3.2 |
| 80 | Teacher | 13 A | 3.8 | 3.4 | 11 E | 3.7 | 3.0 |

TABLE 1—*Continued*

| ITEM | | Per cent Diff. Like | P.E.diff. | Diff./P.E.diff. | Per cent Diff. Dislike | P.E.diff. | Diff./P.E.diff. |
|---|---|---|---|---|---|---|---|
| 85 | Fat men | 12 E | 3.8 | 3.2 | 16 A | 4.1 | 3.9 |
| 86 | Fat women | | | | 21 A | 4.0 | 5.3 |
| 89 | Tall men | | | | 11 A | 3.0 | 3.7 |
| 91 | Short men | | | | 14 A | 3.9 | 3.6 |
| 92 | Short women | 21 E | 4.0 | 5.3 | 19 A | 3.7 | 5.1 |
| 93 | Blondes | 34 E | 3.1 | 11.0 | 8 A | 2.8 | 3.1 |
| 98 | Sideshow freaks | | | | 11 A | 3.2 | 3.4 |
| 101 | People with protruding jaws | | | | 13 A | 4.0 | 3.3 |
| 103 | People with hooked noses | | | | 13 A | 4.1 | 3.2 |
| 110 | Bow-legged people | | | | 13 A | 3.9 | 3.3 |
| 113 | Very polite people | 11 A | 2.9 | 3.8 | | | |
| 120 | People who borrow things | | | | 19 E | 4.0 | 4.8 |
| 122 | Optimists | 14 E | 3.7 | 3.8 | | | |
| 127 | People easily led | | | | 11 A | 4.0 | 2.8 |
| 138 | Irreligious people | | | | 12 A | 3.8 | 3.2 |
| 147 | People who have no sense of humor | | | | 14 E | 3.6 | 3.9 |
| 155 | Methodical people | 13 E | 3.5 | 3.6 | 13 A | 3.3 | 3.9 |
| 156 | Fashionably dressed people | 13 E | 4.1 | 3.2 | 9 A | 2.7 | 3.3 |
| 158 | Southerners | 11 E | 4.1 | 2.7 | 12 A | 2.8 | 4.3 |
| 162 | Golf | 18 E | 3.6 | 5.0 | | | |
| 165 | Tennis | 20 E | 4.0 | 5.0 | | | |
| 172 | Poker | 20 E | 3.3 | 6.1 | 25 A | 3.6 | 6.9 |
| 174 | Billiards | 21 E | 4.5 | 4.7 | 15 A | 3.7 | 4.1 |
| 175 | Dancing | 18 E | 4.6 | 3.9 | 14 A | 3.0 | 4.7 |
| 179 | Excursions | 11 A | 4.5 | 2.4 | 14 E | 1.7 | 8.2 |
| 180 | Smokers | 17 E | 3.6 | 4.7 | 23 A | 4.0 | 5.8 |
| 192 | "Life" | 17 E | 3.3 | 5.2 | | | |
| 201 | William S. Hart | | | | 16 E | 3.0 | 5.3 |
| 209 | Serial movies | | | | 18 E | 4.0 | 4.5 |
| 213 | Telling a story | 9 A | 3.5 | 2.6 | 14 E | 4.2 | 3.3 |
| 215 | Arguments | 12 E | 3.7 | 3.2 | 8 A | 1.9 | 4.2 |
| 221 | Teaching | 16 A | 3.7 | 4.3 | 21 E | 3.8 | 5.5 |
| 224 | Repairing electrical wiring | 26 E | 3.4 | 7.7 | 13 A | 3.0 | 4.3 |

TABLE 1—*Continued*

| ITEM | | Per cent Diff. Like | P.E. diff. | Diff. / P.E. diff. | Per cent Diff. Dislike | P.E. diff. | Diff. / P.E. diff. |
|---|---|---|---|---|---|---|---|
| 225 | Cabinet making | 10 E | 3.8 | 2.6 | 14 A | 3.3 | 4.2 |
| 227 | Digging in the garden | 27 A | 3.2 | 8.4 | 30 E | 3.5 | 8.6 |
| 234 | Continually changing activity | 19 E | 4.0 | 4.8 | 11 A | 3.7 | 3.0 |
| 235 | Living in the city | 30 E | 4.1 | 7.3 | 35 A | 3.6 | 9.7 |
| 237 | Sunday Blue Laws | 18 A | 3.5 | 5.1 | 33 E | 3.8 | 8.7 |
| 238 | Art | 13 E | 3.8 | 3.4 | | | |
| 240 | Chemistry | 30 E | 3.2 | 9.4 | 9 A | 2.9 | 3.1 |
| 247 | Languages, ancient | 22 E | 4.1 | 5.4 | | | |
| 252 | Mechanical drawing | 44 E | 3.1 | 14.4 | 11 A | 2.5 | 4.4 |
| 253 | Music | 11 E | 3.2 | 3.4 | | | |
| 254 | Penmanship | 16 A | 3.1 | 5.1 | 15 E | 3.0 | 5.0 |
| 259 | Shop work | 25 E | 3.9 | 6.4 | | | |
| 262 | Zoology | 23 A | 3.9 | 5.9 | 8 E | 2.9 | 2.8 |
| 263 | Caring for horses | 47 A | 3.6 | 13.1 | 31 E | 3.0 | 10.3 |
| 264 | Beef cattle feeding | 58 A | 3.4 | 17.1 | 38 E | 3.0 | 12.7 |
| 265 | Hog producing | 64 A | 3.3 | 19.4 | 45 E | 2.8 | 16.1 |
| 266 | Purebred livestock production | 61 A | 3.3 | 18.5 | 34 E | 2.8 | 12.1 |
| 267 | Producing for a special trade | 50 A | 3.7 | 13.5 | 33 E | 2.8 | 11.8 |
| 268 | Peddling produce | 13 A | 2.7 | 4.8 | 20 E | 4.2 | 4.8 |
| 269 | Milking | 36 A | 3.6 | 10.0 | 35 E | 3.7 | 9.5 |
| 270 | Raising calves | 54 A | 3.5 | 15.4 | 44 E | 3.2 | 13.7 |
| 271 | Corn growing | 59 A | 3.4 | 17.4 | 44 E | 2.9 | 15.2 |
| 272 | Repairing buildings and fences | 33 A | 3.9 | 8.5 | 30 E | 3.0 | 10.0 |
| 273 | Beautifying the farmstead | 54 A | 3.0 | 18.0 | 24 E | 2.3 | 10.4 |
| 274 | Caring for poultry | 60 A | 3.4 | 17.6 | 38 E | 3.2 | 11.9 |
| 275 | Coöperative marketing | 53 A | 3.6 | 14.7 | 32 E | 2.6 | 12.3 |
| 276 | Market analysis | 43 A | 3.7 | 11.6 | 27 E | 2.6 | 10.4 |
| 277 | Gardening | 47 A | 3.7 | 12.7 | 28 E | 3.2 | 8.8 |
| 278 | Rural sociology | 37 A | 3.6 | 10.3 | 33 E | 2.9 | 11.4 |
| 279 | General farming | 58 A | 3.5 | 16.6 | 37 E | 3.0 | 12.3 |
| 280 | Fruit farming | 54 A | 3.5 | 15.4 | 28 E | 2.8 | 10.0 |
| 281 | Selling feeds | 28 A | 3.6 | 7.8 | 28 E | 3.3 | 8.5 |
| 282 | Land appraising | 12 A | 3.4 | 3.5 | 16 E | 3.5 | 4.6 |

TABLE 1—*Continued*

| ITEM | | Per cent Diff. Like | P.E. diff. | Diff. / P.E. diff. | Per cent Diff. Dislike | P.E. diff. | Diff. / P.E. diff. |
|---|---|---|---|---|---|---|---|
| 283 | County agent | 56 A | 3.5 | 16.0 | 30 E | 3.2 | 9.4 |
| 284 | Extension work in agriculture | 73 A | 2.9 | 25.2 | 39 E | 2.6 | 15.0 |
| 285 | Agricultural research | 52 A | 3.3 | 15.8 | 33 E | 2.8 | 11.8 |
| 286 | Livestock commission | 52 A | 3.6 | 14.4 | 31 E | 2.8 | 11.1 |
| 287 | Poultry marketing | 39 A | 3.6 | 10.8 | 26 E | 3.4 | 7.7 |
| 288 | Farm bureaus work | 55 A | 3.6 | 15.3 | 32 E | 2.9 | 11.0 |
| 289 | Farm demonstration work | 58 A | 3.5 | 16.6 | 31 E | 2.9 | 10.7 |
| 290 | Fertilizer | 51 A | 3.6 | 14.2 | 31 E | 2.9 | 10.7 |
| 291 | Angus cattle | 38 A | 3.6 | 10.6 | 24 E | 3.1 | 7.7 |
| 292 | Jersey cattle | 50 A | 3.7 | 13.5 | 19 E | 2.9 | 6.6 |
| 293 | Trotting horses | 16 A | 4.1 | 3.9 | | | |
| 294 | Draft horses | 51 A | 3.5 | 14.6 | 20 E | 2.5 | 8.0 |
| 295 | Two-row cultivator | 54 A | 3.6 | 15.0 | 25 E | 2.6 | 9.2 |
| 296 | Clover huller | 30 A | 3.7 | 8.1 | 20 E | 2.9 | 6.9 |
| 297 | Four horse team | 49 A | 3.7 | 13.2 | 23 E | 2.7 | 8.5 |
| 298 | Milking machine | 47 A | 3.7 | 12.7 | 22 E | 2.6 | 8.5 |
| 299 | Stock show | 54 A | 3.5 | 15.4 | 25 E | 2.2 | 11.4 |
| 300 | Judging stock | 60 A | 3.5 | 17.2 | 30 E | 2.6 | 11.5 |
| 301 | Corn club | 50 A | 3.5 | 14.3 | 32 E | 2.6 | 12.3 |
| 302 | Club camp | 47 A | 3.7 | 12.7 | 26 E | 2.5 | 10.4 |
| 303 | Club round-up | 54 A | 3.5 | 15.4 | 25 E | 2.4 | 10.4 |
| 304 | Hot-beds | 40 A | 3.6 | 11.1 | 18 E | 3.1 | 5.8 |
| 305 | Cold-frame | 20 A | 3.9 | 5.1 | 16 E | 3.2 | 5.0 |
| 307 | Churning | 21 A | 2.7 | 7.8 | 24 E | 3.3 | 7.3 |
| 311 | Skilled heavy work | 12 A | 3.4 | 3.5 | | | |
| 312 | Slow movements | 12 A | 3.7 | 3.2 | 9 A | 4.2 | 2.1 |
| 328 | Work in one town | 18 E | 3.8 | 4.7 | | | |
| 335 | Work in an office | | | | 22 A | 4.0 | 5.5 |
| 336 | Work requiring teaching | 15 A | 4.1 | 3.7 | | | |
| 337 | Welfare work | 20 A | 3.5 | 5.7 | 18 E | 3.5 | 5.1 |
| 338 | Taking part in entertainments | | | | 39 E | 4.1 | 9.5 |
| 341 | Increasing the amount of production | 12 A | 3.7 | 3.2 | | | |
| 344 | Less pay with safety | 19 A | 4.1 | 4.6 | | | |
| 345 | Work with children | 28 A | 3.9 | 7.2 | 18 E | 3.1 | 5.8 |
| 348 | Laboratory work | 20 E | 3.8 | 5.3 | | | |

The differences of these percentages were then determined, and their probable errors calculated from the formula[2]

$$\text{P.E.}_{\text{diff.}} = \sqrt{\text{P.E.}_{\text{I}}^2 + \text{P.E.}_{\text{II}}^2}$$

Usually the items which yielded a ratio of the difference to its probable error greater than 3 for either "like" or "dislike" were selected as indicative of differential vocational interests for the two groups in question. A few items falling short of this criterion were, however, included. The items selected are listed with their percentage differences and the probable errors of these differences in table 1. The numbers before the items indicate the order of the items in the blank as originally administered. The letter E or A after a percentage difference means that students of engineering or of agriculture indicated a greater liking or dislike for the item in question. Table 1 is to be read, then: The item "Actor" was marked as liked by 14 per cent more engineering students than agricultural students. The P.E. of this percentage is 3.4, and the ratio of the difference to its P.E. is 4.1. The item was marked as disliked by 8 per cent more agricultural students than engineering students. The P.E. of this percentage is 3.9, and the ratio of the difference to its P.E. is 2.0.

The items listed in table 1 were then used as a basis for scoring approximately a hundred randomly selected papers each for students of agriculture and for students of engineering. The items indicating engineering interests were marked plus, and those indicating agricultural interests were marked minus. The total score was then obtained by giving each item a weighting of one and summing the pluses and minuses.

## IV. THE DATA

The basic data on the various items and their significance as differential criteria are given in table 1.

The results for the papers scored according to the method

[2] Yule, G. U. *Op. cit.*, p. 269, formula (6).

previously described were then tabulated for the purpose of studying the group discrimination achieved by this method. Table 2 gives the frequency distributions obtained for students

TABLE 2

*Frequency distributions showing the discriminatory power of the Purdue Interest Report Blank as between students of agriculture and engineering*

| SCORE | AGRICULTURE FRESHMEN | ENGINEERING FRESHMEN | AGRICULTURE SENIORS |
|---|---|---|---|
| −44 to −47 | 1 | | 2 |
| −40 to −43 | 5 | | |
| −36 to −39 | 8 | | 2 |
| −32 to −35 | 12 | | |
| −28 to −31 | 14 | | 2 |
| −24 to −27 | 9 | | 1 |
| −20 to −23 | 19 | 1 | 2 |
| −16 to −19 | 6 | | 5 |
| −12 to −15 | 14 | 1 | |
| −8 to −11 | 6 | 2 | 5 |
| −4 to −7 | 6 | 3 | 2 |
| 0 to −3 | 2 | 4 | 1 |
| +1 to 4 | | 8 | |
| +5 to 8 | 1 | 11 | |
| +9 to 12 | 1 | 14 | 1 |
| +13 to 16 | | 21 | 1 |
| +17 to 20 | | 18 | |
| +21 to 24 | | 14 | |
| +25 to 28 | | 4 | |
| +29 to 32 | | 5 | |
| Total..... | 112 | 106 | 23 |
| Median..... | −22.53 | +14.33 | −17.2 |
| Mean..... | 21.25 | +13.44 | |

Biserial eta = 0.95

of agriculture and engineering. Figure 1 is a graphic representation of these data. It is obvious that the instrument is on the whole quite successful in separating two such groups selected at random within their respective vocational schools in the uni-

versity. The overlapping is very slight. So far as the meager data on agricultural seniors warrant any inference, they indicate relative stability of interests during four years of college.

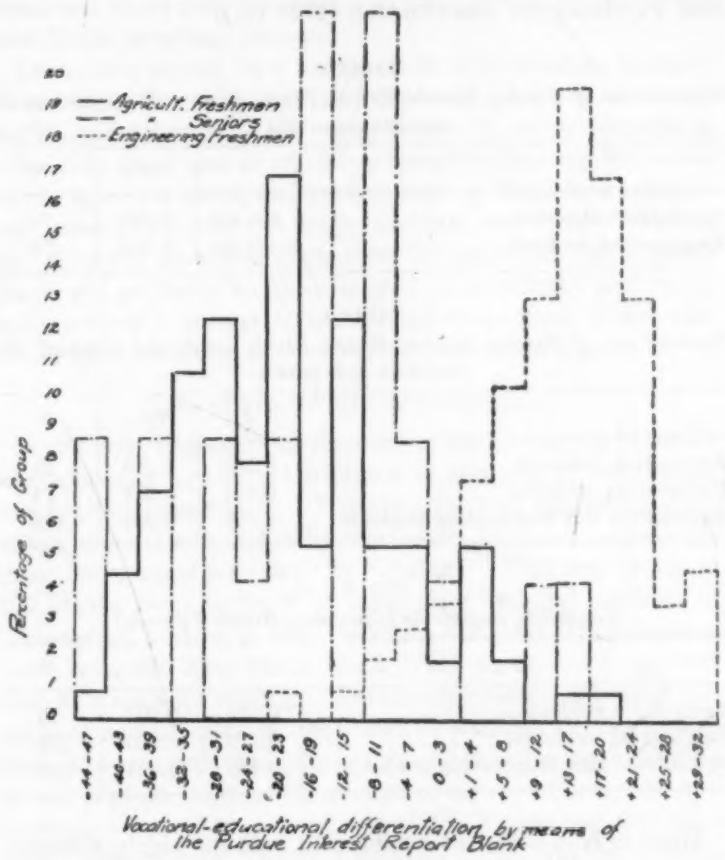



FIG. 1

It would seem, then, that in so far as this separation into two groups is a valid criterion, the revised blank is a valid instrument in determining vocational interests of the sort in question.

A sort of negative criterion is afforded by the lack of correlation between mental-scholastic tests and scores on the Purdue Interest Report Blank. Table 3 gives the relevant data for the $r$'s of interest scores with the National Council on Education Psychological Examination (1927 ed.).

TABLE 3

*Correlations of Purdue Interest Report Blank scores with psychological examination scores*

| | $r$ | P.E.$_r$ | N |
|---|---|---|---|
| Agriculture students | +0.419 | 0.063 | 78 |
| Engineering students | +0.188 | 0.076 | 74 |

TABLE 4

*Correlations of Purdue Interest Report Blank scores and sums of all freshman year grades*

| | $r$ | P.E.$_r$ | N |
|---|---|---|---|
| Agriculture students | 0.036 | 0.076 | 78 |
| Engineering students | 0.061 | 0.078 | 74 |
| Agriculture and Engineering students | 0.018 | 0.054 | 152 |

TABLE 5

*Reliability coefficients (Spearman Brown Formula)*

| | $r$ | P.E.$_r$ | N |
|---|---|---|---|
| Agriculture students | 0.810 | 0.024 | 95 |
| Engineering students | 0.678 | 0.036 | 100 |
| Agriculture and Engineering students | 0.838 | 0.014 | 195 |

There is here sufficient evidence to indicate widely different functions measured by the two different instruments.

Correlations were also worked out with the scholastic grades achieved during the freshman year. They are shown in table 4.

Whatever is measured by the interest blank, it is not scholastic ability as required in the first year of academic work.

That this lack of correlation is not a result of the unreliability of the Purdue Interest Report Blank is shown by the fact that the coefficient of reliability is fairly high. Table 5 indicates the reliabilities obtained by correlating the odd *vs.* the even numbered items and "stepping up" these values by the Spearman-Broan prophecy formula.

That there should be a notable lack of correlation between scores on the interests blank and academic grades is not necessarily condemnatory of the instrument. It seems plausible *a priori* that there should be little connection between the scores on the interests blank and performance in English, mathematics, history, chemistry, biology, or what not. Such interests and attitudes as reveal themselves in scores on the interests blank will probably be more related to specifically and much more narrowly defined vocational activities than is the case with such subjects as are studied in the freshman year.

## V. SUMMARY AND CONCLUSIONS

A differential diagnostic instrument for measuring interests of students of agriculture *vs.* those of students of engineering was built by selecting out of an item-by-item tabulation those which showed a statistically significant difference as between these two groups for "like" or "dislike." This was based on 100 blanks from students of agriculture and 200 students of engineering, making a total tabulation of 104,700 responses, there being 349 items in the blank. The significant items were used as a basis for scoring 112 randomly selected blanks of students of agriculture and 106 blanks of students of engineering, each item being marked plus or minus depending upon whether it indicated an engineering interest or an agricultural interest. Each item was given a weighting of 1. A high negative score would thus indicate a large number of agricultural interests, a high positive score, a large number of engineering interests.

The following conclusions seem warranted on the basis of the data presented.

1. It is possible to measure with a fair degree of reliability the divergence of interests as between students who enroll in agriculture and those who enroll in engineering.

2. The divergence of the two samplings studied was quite pronounced. No score in either group overlaps the mean of the other. Bi-serial eta yielded a value of 0.95.

3. The correlation of whatever is measured by the Purdue Interest Report Blank with mental-scholastic aptitude tests or with scholastic achievement is negligible.

4. Should further investigation prove these interests as measured to hold for those who are successful in the types of vocations here in question, a valuable instrument for aid in vocational-educational counseling is indicated.

## REFERENCES

(1) Brandenburg, G. C., "Analyzing Personality." Jour. Applied Psychol., June and September, 1925.

(2) Bridges, J. M., and Dollinger, V. M., "The Correlation between Interests and Abilities in College Courses." Psychol. Review, vol. 27, 308–14, July, 1920.

(3) Cowdery, K. M., "Measurement of Professional Attitudes." Jour. Personnel Research, August, 1926.

(4) Crathorne, A. R., "Change of Mind between High School and College as to Life Work." School and Society, Vol. 11, 28–30.

(5) Freyd, Max, "Measurement of Interests in Vocational Selection." Jour. Personnel Research, 1923, vol. 1, 319–328.

(6) Freyd, Max, "A Method for the Study of Vocational Interests." Jour. Applied Psychol., 1922, vol. 6, 243–252.

(7) Franklin, E. E., "The Permanence of Vocational Interests of Junior High School Pupils." Johns Hopkins University Studies in Education, No. 8.

(8) Fryer, Douglas, "Vocational Self-Guidance." 1925, 385 pp.

(9) Fryer, Douglas, "Interest and Ability in Educational Guidance." Jour. Educ. Research, June, 1927, vol. 16, 26–39.

(10) Hubbard, Ruth M., "Interests Studied Quantitatively." Jour. Personnel Research, vol. 4, 365–378, January–February, 1926.

(11) Kitson, H. P., "Interests in Vocational Guidance." Educ. Review, vol. 52, 349–356.

(12) Kitson, H. P., "Economic Implications in the Psychological Doctrine of Interests." Jour. Political Economy, vol. 28, 332–338.

(13) King, Irving, and Adelstein, Morris, "The Permanence of Interests and Their Relation to Abilities." School and Society, vol. 6, 359–360, September 22, 1917.

(14) Miner, J. B., "A Method for Evaluating a Psychograph for Vocational Guidance." Jour. Educ. Psychol., May, 1926, pp. 331–340.

(15) MINER, J. B., "An Aid to the Analysis of Vocational Interests." Jour. Educ. Research, vol. 5, 311–323.

(16) MOORE, B. V., "Some Principles and Practices of Personnel Selection with Special Reference to Graduate Engineers." Psychol. Monograph, vol. 30, 1921.

(17) MOORE, B. V., "A Tested Method of Using Tests for Vocational Guidance." School and Society, 1923, vol. 18, 761–764.

(18) PROCTOR, WILLIAM M., "Educational and Vocational Guidance." Chapter dealing with "Relation of Special Interests and Abilities to School and Vocational Success." 1923.

(19) REAM, M. J., "Ability to Sell." Baltimore, 1924, The Williams & Wilkins Company.

(20) STRONG, E. K., JR., "An Interest Test for Personnel Managers." Jour. Personnel Research, September, 1926.

(21) STRONG, E. K., JR., "Interest Analysis of Personnel Managers." Jour. Personnel Research, October, 1926.

(22) STRONG, E. K., JR., "Differentiation of Certified Public Accountants from other Occupational Groups." Jour. Educ. Psychol., 1927, vol. 18, 227–237.

(23) THORNDIKE, E. L., "The Permanence of Interests and Relation to Abilities." Pop. Sci. Monthly, vol. 81, 449–456, 1922.

(24) THORNDIKE, E. L., "Early Interests: Their Permanence and Relation to Abilities." School and Society, vol. 5, 178–79.

(25) THORNDIKE, E. L., "The Correlation Between Interests and Abilities in College Courses." Psychol. Review, vol. 28, 374–376, 1921.

# STUDIES OF TYPOGRAPHICAL FACTORS INFLUENCING SPEED OF READING

## II. Size of Type

DONALD G. PATERSON and MILES A. TINKER

*University of Minnesota*

Demonstration of the suitability of the Chapman-Cook Speed of Reading Tests (Form A and Form B) in studying the influence of type form on reading performance[1] opens up fruitful lines of attack upon a number of problems in the psychology of reading.

Ideal conditions for determining the influence of various factors such as type form, or size of type on speed of reading should be kept in mind and the extent to which our present technique meets these requirements should be set forth at the outset of our discussion.

### METHODOLOGICAL REQUIREMENTS

1. A sufficiently large number of cases should be tested to insure the statistical validity of obtained means and variabilities entering into the statistical comparisons. The group testing method as here employed makes it easy to fulfill this requirement.

2. The difficulty of the text material should remain constant from study to study and should be readily available to all investigators for repetition and verification of obtained results. The use of standardized speed of reading tests meets this requirement.[2]

[1] Miles A. Tinker and Donald G. Paterson, Influence of Type Form on Speed of Reading, J. Appl. Psychol., Vol. XII, No. 4, August, 1928, pp. 359–368.

[2] The Chapman-Cook Speed of Reading Tests, Forms A and B, may be procured from J. B. Lippincott Co., Philadelphia, Pa. The writers wish to acknowledge their indebtedness to the publishers for permission to reprint these tests for experimental purposes.

3. Alternate forms of textual material should be arranged so that the form used for the second trial is sufficiently more difficult to just offset whatever practice effects may be present. It so happens that Forms A and B of the Chapman-Cook Speed of Reading Test are nicely adjusted in this respect. Data substantiating this statement were obtained from the administration of this test to 2260 high school seniors in Minneapolis and St. Paul during the spring of 1928 (just prior to graduation). The mean score was 17.465 and the probable error of the mean was 0.068 for Form A. The mean score for Form B (which followed Form A in the administration of the tests) was 17.479 with a probable error of the mean of 0.063. In our previous paper this same fact of equality was disclosed in spite of radically different type forms used in printing Forms A and B.[3] Therefore, there is ample justification for assuming that the relative difficulty of Forms A and B and the practice effects in passing from trial one to trial two will remain constant for the level of ability represented by high school seniors and college sophomores even though the type in Form B varies from the type in Form A in size or in any other respect. Any differences in speed of reading arising as a result of such variation in the printing of Form B would be direct evidence of the effect of such change.

4. A testing technique should be employed which will provide an *objective* check on comprehension so that the total amount read within a given time limit may be ascertained without error. The Chapman-Cook technique consists in presenting a series of paragraphs in each of which one word spoils the meaning of the paragraph. This wrong word always occurs in the second half of each paragraph and is wrong because it is out of harmony with the meaning of the paragraph as a whole but is not out of harmony with the meaning of the second half of the paragraph. For example:

> When I am enjoying anything very much, time seems to go very quickly. I noticed this the other day, when

[3] See Tinker and Paterson, op. cit., table 1, p. 362.

I spent the whole afternoon reading a very uninteresting book.

One cannot underline the incorrect word in each paragraph without actually comprehending the whole paragraph, hence the underlinings serve as an ingenious device for objectively checking comprehension.

5. A testing technique which will yield a rate score with accuracy as a negligible factor should be used in order that complicated and uncertain statistical devices for combining the troublesome variables of speed and accuracy into a single index may be dispensed with. The Chapman-Cook test meets this requirement because the vocabulary content is simple, hence comprehension as a factor plays a minor rôle. At least this is proven for college sophomores by the fact that 99.7 per cent of all the paragraphs are marked correctly by them.[4]

6. Test forms should be long enough to permit a sampling of each individual's speed of reading performance so that the resulting scores may be sufficiently reliable in the technical sense of "test reliability" to warrant the use of those scores in group comparisons. This last requirement would be met if the reliability of the test could be shown to be greater than $+0.40$.[5] The Chapman-Cook test more than meets this requirement also. The correlation between the paired scores for Form A and Form B for the 2260 Twin City high school seniors was $+0.76 \pm 0.006$. The correlation between the paired scores for Form A and Form B in sub-groups of 80 students each, as reported later in this paper, ranged from $+0.75 \pm 0.03$ to $+0.83 \pm 0.02$.

[4] See Tinker and Paterson, op. cit., p. 362.

[5] W. S. Hunter in an article entitled a Reply to Professor Carr on "The Reliability of the Maze Experiment" in the J. of Comp. Psychol., 1926, 6, pp. 393–398 quotes a letter from Truman Lee Kelley as follows: "As a practical matter dealing with school grade populations and such differences in reading, arithmetic, spelling, etc., as are commonly found, I have advised my students not to attempt a study of individual differences unless the reliability was 0.90, or better, and not to attempt a study of group differences (assuming classes of about 30) unless the reliability was 0.40, or better."

Search of the literature on the influence of size of type, or of type form on speed of reading failed to reveal any comprehensive studies utilizing techniques and controls which meet the requirements as outlined above. It is this fact that justifies emphasis upon our method of procedure as well as upon the results we have obtained.

### PREVIOUS WORK ON INFLUENCE OF SIZE OF TYPE

Blackhurst[6] measured the rate and accuracy of the reading of pupils in the first four grades with type varying in size as follows: 8, 10, 12, 14, 18, 24, and 30 point. The results indicated that 18 point type is optimal in the third and fourth grades and 24 point type is the optimum for the first and second grades. It is to be noted that striking differences were not obtained although if a method for combining rate and accuracy had been used the differences might have been more significant. This study is not directly comparable to our investigation because he studied oral reading with young children and used the same text material for the various sizes of type attempting to neutralize the effect of practice and learning by systematic variation in presentations. Furthermore, Blackhurst fails to publish the exact text used or to specify the length of line employed.

Even Gilliland's[7] study with adults is not directly comparable with the study reported in this article because he varied the length of line along with variation in the size of type. He printed as a standard a short paragraph of only 12 lines totalling 100 words in 12 point type, with a line length of $3\frac{1}{3}$ inches, and then by photographic reduction and enlargement changed size of type (*and length of line*) to provide samples of 3, 4, 6, 9, 12, 18, 36, 54, 72, and 90 point type. Twenty-four adult subjects were used in one experiment and from 23 to 52 subjects

[6] J. H. Blackhurst, Investigations in the Hygiene of Reading, Pub. by Warwick & York, Inc., Baltimore, 1927, pp. 41–47.

[7] A. R. Gilliland, The Effect on Reading of Changes in the Size of Type. El. Sch. J., 1923, 24, 138–146.

in the various comparisons in another experiment. Each read reduced or enlarged sample and then re-read the same material in the 12 point type standard, continuing in this manner until all samples were read. Gilliland concludes, "The reading of the average adult is not greatly affected by changes in size of type between the limits of 36 point and 6 point type. Above 36 point type or below 6 point type the rate of reading begins to decrease for a majority of the subjects." The failure to show differences between the rather large limits of 6 point and 18 point type may have been due to inadequate experimental techniques since there was a failure to control practice effects. Furthermore, such a small amount of text must have resulted in very low reliabilities for the single measurements (the average time of reading the paragraph was in the neighborhood of seven seconds) and if they were below +0.40 the obtained group comparisons would be invalidated. The reported average reading time of about eight seconds per 100 words yields a rate of about twelve words per second which is more than double the average reported by most investigators. This suggests failure to provide adequate checks on comprehension. Absence of significant differences for such large variations in size of type may also have been due to the unfortunate selection of a size of type and a line length as a standard which happens not to be the optimum.

Griffing and Franz,[8] using twelve college students, compared the speed of reading identical selections from the Bible printed in 12 point and in 5 point type. Their results showed that twelve point type was read 10 per cent faster. Length of line, however, was not specified.

Authors in the field of the psychology of advertising would probably agree as to the importance of size of type on speed of reading but owing to absence of experimental data either fail to mention the problem or proceed to discuss it in rather general terms. Starch points out that type smaller than 10 point becomes increasingly difficult to read as it decreases in

[8] H. Griffing and S. I. Franz. On the Conditions of Fatigue in Reading. Psychol. Rev., 1896, III, 513-530.

size.[9] By use of the historical method he shows that advertisers themselves have arrived at the same conclusion. Whereas 85 per cent of full-page advertisements in the *Century Magazine* had the main body text printed in 8 point type or smaller in 1881 only 5 per cent of such advertisements in the same magazine followed this practice in 1910. Hollingworth confines his remarks on the importance of size of type to Rothlein's work reported in the *American Journal of Psychology* for January, 1912, which is confined to legibility of single letters in isolation and in non-sense syllables.[10] Scott, Adams, and Poffenberger in their texts on advertising fail to mention the problem.[11]

This survey of the literature indicates the need of further study of this problem in which length of line is held constant while size of type alone is varied together with adequate controls regarding practice effects, length of text, number of subjects, and checks on comprehension.

## PROCEDURE

In our previous paper dealing with lower case, all capitals, and italic type forms, we were forced to use an A B B A sequence because the combined influence of practice and difficulty of the two forms was unknown. The results of that investigation revealed the approximate equality of the two forms when Form B follows Form A. The testing of high school seniors, previously mentioned, likewise demonstrated the precision with which the increased difficulty of Form B equals the practice effect in passing from trial one to trial two. In our further work, therefore, the A B B A sequence method was abandoned permitting a simpler and more straightforward comparison be-

[9] D. Starch. Principles of advertising. Pub. by A. W. Shaw Co., Chicago, 1923, p. 663.

[10] H. L. Hollingworth, Advertising and Selling, pub. by D. Appleton & Company, N. Y., 1920, pp. 185–187.

[11] W. D. Scott, The Theory of Advertising, pub. by Small, Maynard & Co., Boston, 1903; H. F. Adams, Advertising and Its Mental Laws, Macmillan Co., N. Y., 1916; A. T. Poffenberger, Psychology in Advertising, A. W. Shaw Co., Chicago, 1925.

tween reading performance on Form A set up in lower case, 10 point old style, 80 mm. lines of print, as the *standard* and on Form B set up in lower case letters differing with respect to size of type as the *variable* isolated for study.

In all, 320 students in sophomore general psychology were tested. None of these students had previously been exposed to the Chapman-Cook Speed of Reading Tests. The subjects were selected at random from two lecture sections totalling about 400 and 500 students respectively. This yielded the following test groups of 80 students each:

Group I. Form A, 10 point, Form B, 6 point
Group II. Form A, 10 point, Form B, 8 point
Group III. Form A, 10 point, Form B, 12 point
Group IV. Form A, 10 point, Form B, 14 point

The same testing procedure as was described in our previous report on the influence of type form was followed except that the tests were given to the two large lecture sections by one of the writers instead of being administered to some twenty-five small recitation sections. As in the previous study, coöperation was secured by emphasizing the importance of speed of reading in college work and by promising to inform each student regarding his standing in comparison with sophomore averages. The real purpose of the experiment was not explained and each student presumably believed that all other students were being given the same two set-ups. The time limit was $1\frac{3}{4}$ minutes.

Scoring was greatly facilitated by merely noting the total number of paragraphs marked after making sure in each case that the directions were fully understood. This simplified scoring is justified because our previous study showed that college sophomores maintain a 99.7 per cent accuracy in performing this test.

## RESULTS

Table 1 shows that in every comparison the 10 point standard was read more rapidly than any other size of type. For ex-

ample, in Group I, the third column shows that the average number of paragraphs read was 18.81 for Form A, 10 point and was 17.71 for Form B, 6 point. In the fourth and fifth columns are given the probable error of the distribution and the probable error of the mean for each form. In the sixth,

TABLE 1

*Influence of size of type*

Differences given are for the mean score on Form A, 10 point, 80 mm. line minus the mean score on Form B, varying in type size (6, 8, 12, or 14 point), 80 mm. line. The mean score is the average number of paragraphs of 30 words each in the Chapman-Cook Speed of Reading Tests read in 1¼ minutes. In each sub-group, N = 80 students in sophomore general psychology.

| TEST GROUP | FORM AND SIZE | MEAN | P.E. dist. | P.E.M | DIFFERENCES BETWEEN MEANS IN | | | P.E. diff. | $\frac{D}{P.E.\text{diff.}}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Paragraphs | Words per second | Per cent | | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| I | A, 10 point | 18.81 | 2.96 | 0.33 | 1.10 | 0.31 | 6.2 | 0.19 | 5.64 |
| | B, 6 point | 17.71 | 2.61 | 0.29 | | | | | |
| II | A, 10 point | 19.30 | 2.97 | 0.33 | 0.95 | 0.27 | 5.2 | 0.18 | 5.13 |
| | B, 8 point | 18.35 | 2.51 | 0.28 | | | | | |
| III | A, 10 point | 18.93 | 3.17 | 0.35 | 1.04 | 0.30 | 5.8 | 0.20 | 5.25 |
| | B, 12 point | 17.89 | 2.53 | 0.28 | | | | | |
| IV | A, 10 point | 18.29 | 2.85 | 0.32 | 1.18 | 0.34 | 6.9 | 0.20 | 5.94 |
| | B, 14 point | 17.11 | 2.45 | 0.27 | | | | | |

seventh, and eighth columns the differences between the averages for Forms A and B are given in terms of paragraphs read, words read per second, and per cent increase in rate of reading Form A as compared with Form B. In the ninth column will be found the probable error of the obtained differences between

the means given in column 3.[12] Finally, in the tenth column, the ratio of the obtained difference to the probable error of the difference is shown, i.e., the value in column 6 divided by the value in column 9.

Column 6 of the table shows that the obtained differences are consistently in the same direction, i.e., 6, 8, 12, and 14 point type are all read slower than 10 point type. The differences in number of words read per second appear small (0.27 to 0.34 words per second; see column 7) yet when these are expressed as percentages (column 8) we find that the differences range from 5.2 to 6.9 per cent and thus become quite meaningful. The ratios in the last column of the table indicate that all the obtained differences are highly reliable.

It is evident from the results contained in table 1, that, for the sizes of type here investigated (6, 8, 10, 12, and 14 point) and with line length held constant at 80 mm., 10 point type yields the fastest reading and is thus demonstrated to be the optimum size.

From the viewpoint of economy of space 6 point type is most economical. In comparing the amount of space taken for the different texts we find that 8 point took 24 per cent; 10 point,

[12] The values for P.E.$_{\text{diff.}}$ shown in column 9 were obtained by use of the formula for *correlated measures*, i.e.,

$$\text{P.E.}_{\text{diff.}} = \sqrt{\text{P.E.}^2_{M_1} + \text{P.E.}^2_{M_2} - 2r\text{P.E.}_{M_1}\text{P.E.}_{M_2}}$$

(See Garrett, H. E., Statistics in Psychology and Education, pub. by Longmans, Green & Co., N. Y., 1926, pp. 286-288.) The correlations for the four sub-groups in this experiment were $+0.81 \pm 0.03$ for Group I; $+0.83 \pm 0.02$ for Group II; $+0.83 \pm 0.02$ for Group III; and $+0.75 \pm 0.03$ for Group IV. It is to be noted that this formula increases the statistical significance of the obtained differences as compared with the formula for uncorrelated measures, i.e., $\text{P.E.}_{\text{Diff.}} = \sqrt{\text{P.E.}^2_{M_1} + \text{P.E.}^2_{M_2}}$. We used this latter formula erroneously in our previous paper on type form, therefore the $\frac{D}{\text{P.E.}_{\text{diff.}}}$ value reported on p. 364 should be 21.47 instead of 10.1 and the corresponding value reported on p. 365 should be 4.47 instead of 2.23. Blackhurst, op. cit. errs in the same way in reporting his $\frac{D}{\sigma_{\text{diff.}}}$ values.

70 per cent; 12 point, 130 per cent; and 14 point, 219 per cent more space than 6 point type. In variation from our optimum (the ten point type) it was discovered that the 6 point text took 41 per cent less space; the 8 point, 29 per cent less; the 12 point, 37 per cent more; and the 14 point, 87 per cent more space. One is seldom justified in reducing printing space to such a degree that speed of reading is markedly decreased. There would be even less justification for increasing the amount of space (due to size of type) above the amount which yields the fastest reading.

For the larger types (12 point and 14 point) the increased amount of *horizontal* space covered *per word* is probably the most important factor tending to reduce the speed of reading.[13] With the smaller types, however, (6 and 8 point), it is likely that reduction in legibility of the printed characters accompanying reduction in the size of the letters is the important factor which slows up the reading.

## SUMMARY AND CONCLUSIONS

1. The present experiment continues previous work of the writers on mechanical factors influencing speed of reading and is a further demonstration of the suitability of the Chapman-Cook Speed of Reading Tests whereby one mechanical factor at a time may be varied coupled with adequate controls as regards practice effects, length of text, number of subjects, and checks on comprehension. The literature on the effect of size of type on speed of reading reveals the need of just such improvements in methodology if results of greater universal validity are to be secured.

2. A speed of reading record was obtained for texts set up in 6, 8, 10, 12, and 14 point type with line length kept constant at 80 mm. Comparisons were made between 10 point (as a standard) and each of the other sizes of type. There were 320 college students used in the experiment.

[13] See M. A. Tinker, Numerals Versus Words for Efficiency in Reading. Jour. Appl. Psychol., 1928, 12, 190–199.

3. Texts in 6, 8, 12, and 14 point type were read more slowly than the text in 10 point. The differences ranged from 5.2 to 6.9 per cent and were all statistically significant. The increasingly important rôle of reading in all phases of modern life in which every possible scientific application and every feasible mechanical invention is promptly utilized in the race for efficiency emphasizes the practical significance of differences in speed of reading as small as 5 per cent. In the total amount of reading done in the lifetime of millions of people a saving of 5 per cent bulks extremely large.

4. The results obtained in this experiment justify the conclusion that, for a line length of 80 mm. and for the comparisons here made, 10 point type yields the fastest reading and is thus the optimum size of type (in comparison with other sizes used) for efficient reading.

# THE SOCIAL PSYCHOLOGY OF THE AUSTRALIAN ABORIGINAL

S. D. PORTEUS

*University of Hawaii*

When writers wish to instance the lowest depths of primitiveness in humankind they usually choose the Australian aborigine or as they sometimes erroneously call him, the Australian bushman. The latter, by the way, is not an aboriginal at all, but a white man who is familiar with the out-back country. The equivalent term in America would perhaps be frontiersman. The error is due to confusing the Bushman of South Africa with the Australian. As an example of this confusion an editorial writer wishing to express the extremes of polarity said that there was as much difference between two things as there is between a Harvard professor and an Australian bushman. As a matter of fact there might not be such a great difference and in some cases the bushman might quite possibly be the better man.

Leaving aside this error of description—from which by the way, even anthropologists are not exempt—it is proper to enquire whether the Australian is really as primitive and degraded in point of social evolution as he is generally supposed to be. The writer's sincere opinion is that he is not by any means "primitive" but highly specialized and that his social organization has been developed in a most remarkable way to suit the physical conditions in which he found himself. We must remember in all of our discussion on this topic that the Australian arrived in his isolated continent at a period before any animals were domesticated and before the products of agriculture had been developed—or at least he brought none with him. The aborigine came practically empty-handed into

Australia and with his bare hands he has laboured to ensure his survival in that naturally inhospitable continent. Only those who have lived in Australia can realize how difficult human sustenance would be to a people deprived of the means and products of agriculture. Outside of some sparsely scattered nardoo bushes, a few yams, lily roots, a native peach and a nut bearing tree which grows in one small area, there is not a plant which can provide a staple article of diet. With the exception of nardoo and grass seed cakes scarcely any food is storable. The aboriginal eats some part of no less than 300 plants but this wide range is evidence of scarcity rather than the reverse. His larder must be eked out with snakes, lizards, grubs, beetles, insects and when obtainable, the larger marsupials. In the interior water is so scarce that he must sometimes depend for his supply on dew, or liquid stored up in the roots of certain trees. Hence the conditions of life are so extremely difficult that a very special form of social organization has been forced upon the Australian race.

The main features of this organization are an absence of chieftainship, government by the old men of the tribe being the general rule; a division of almost all tribes into totemic clans, associated with which is a system of food restrictions; another division of the tribes into classes so that an individual may not, under penalty of death, marry with women outside the appropriate exogamous group; a system either of male or female descent by which the children of a marriage may be credited to moieties of the tribe other than those to which their parents belong; initiation ceremonies which are in many tribes exceedingly severe; and an almost universal rule of mother-in-law avoidance. These are the customs which are more or less general, and in each case I propose to try to show that the custom itself is either the result of very careful planning or is a specialization very directly advantageous to the tribe under the conditions under which it lives.

Taking the matter of tribal government first, we see that it is the form which would be most naturally present in Australia. So inhospitable to man is this country, that in the whole of the

continent there were probably never more than 200,000 natives living at the one time. The population was divided into many tribes, with perhaps 250 members in each, and the tribe was again divided into groups each having its own hunting territory the bounds of which were known to all its members. Under the circumstances there was very little contact of tribe with tribe and practically nothing in the way of tribal warfare. There was little time nor opportunity for hostile expeditions and in fact there being no moveable wealth except a few weapons, dilly bags, etc., there was really nothing to fight over. There would be no material advantage in the domination of tribe over tribe. An aboriginal Napoleon or Alexander would be wasted in Australia—there would be no worlds at all for him to conquer. Fighting there is in plenty but it is in the nature of individual feuds and forays.

Under these circumstances there were no stories of tribal exploits, no traditions of successful wars, no man who had claims to be the champion or defender of his tribe. Individual reputations died with their owners. Hence there is almost nothing approaching an individual hereditary chieftainship—there are no chiefs and there is nothing to inherit. The struggle for existence is not internecine but against the niggardliness of Nature itself. The men of authority, of prestige, of standing in the tribe will therefore not be its martial heroes but those who by a mere survival to old age have proved their fitness to cope with drought, starvation and the other vicissitudes of life in that most uncertain country. With that all-pervading fear of magic which the Australian possesses to contend against also, the men with knowledge of the appropriate rites and ceremonies will have great community value for they are the repositories of tribal tradition. Hence government will be entirely in the hands of the tribal council, admission to which is rigidly restricted to the mature and fully initiated members of the tribe. Power will be in the hands of the well-instructed rather than of the physically well-endowed. This power is real and far reaching and matters of life and death are within the scope of its operations. Respect is not accorded to the aged by

courtesy but as a matter of right. As Roth[1] says "Old men in general get the best of everything and are treated with respect . . . . They not only get respect but they command fear and obedience." Nor is this old man government confined to a few tribes but is general all over Australia, only one or two tribes having anything approaching hereditary chieftainship and they inhabit the more thickly populated south east where rainfall is more plentiful and sustenance easier.

Authority is naturally expressed in privilege and in a country so inhospitable, the first privileges will naturally pertain to food. Consequently there are certain food restrictions which apply directly to the old men's benefit. The choicest parts of turkey and emu, and such delicacies as black duck, female opossum and honey are in some tribes forbidden to men under forty years of age. Many things are forbidden to women also and the prohibition in their case is usually permanent. All kinds of dire penalties are supposed to attend the infraction of these rules, including sickness, death, or strangely enough, premature old age.

The next step in social organization was probably the devising of a system of totem kins by which each group was made responsible for the carrying out of the rites and ceremonies which insure a plentiful supply of the animal in question. Generally speaking the members of the group cannot eat of their totemic animal. The double effect of this system is to reinforce the other food restrictions by focussing attention on the products of the chase and secondly to ensure tribal unity by making every hunt a coöperative affair since the groups become mutually dependent on each other for part of their sustenance. If, for example, a kangaroo should be killed by men of this totem they must distribute their game among the rest of the tribe, this enforced self-denial being compensated for by the exchange of food forbidden to other totems. The elimination of selfishness is so thorough that rewards made by the white man to a native for some laborious or dangerous task are quite

[1] Roth: North Queensland Ethnography, Bull. 8, p. 7.

naturally enjoyed by relatives or friends who have done absolutely nothing to merit them.

The totemic food restrictions thus have the effect of a very ingenious system of self denial affecting all the members of the tribe in different particulars. Food is too scarce for the tribe to make any animal absolutely taboo, but by spreading the taboos among the various totems, the effects of self denial are obtained without the tribe as a whole suffering to the slightest degree. Other restrictions which effectively disbar selfishness are also prevalent. In some tribes the successful hunter is almost unlucky because all the best parts of his kill must be given to others in a definite order and amount. Food restrictions are thus seen to be one of the most astute devices possible for the preservation of tribal unity. This point has, as far as I know, never been stressed by observers.

After food the next important thing in savage life is women, and our theory is that the old men, having seen that restrictions could be applied to their advantage in one case, devised a similar system for the other.[2] Accordingly the tribes were divided into classes each one having one other class with which the members could intermarry. Some tribes had eight classes, some four and others two. The immediate effect of this division was to limit the number of eligible women for each member of the tribe. In an eight-class system such as that of the Arunta as described by Spencer and Gillen, only one-eighth of the women in the tribe were eligible, as regards marriage, to any man of the tribe. Various theories have been advanced to account for the system. Among others fear of incest was one of those most strongly supported, but as Spencer points out, the evil effects of incest are most unlikely to be realized by individuals, some of whom in certain tribes, see no connection between cohabitation and pregnancy. In certain Poly-

[2] Authorities such as Howitt and Spencer and Gillen agree in thinking that the totemic divisions existed before the exogamous class system, although they do not relate them closely. The classes in many cases have similar names to the totems—usually names of animals and birds. This indicates a relation between the two systems.

nesian peoples brother-sister marriages among those of highest rank were in vogue and to suggest that a sense of tribal morality in this regard was absent in them, but present in the Australians, seems unnecessary. The point seems to have been almost entirely overlooked that while the exogamous class system limits the number of marriageable women it also limits competition, especially for the old men who naturally play the largest part in arranging marriages. Marriage indeed takes the form of barter, a woman being given in marriage in exchange for a tribal sister of the husband. Seeing that the moveable wealth of the aboriginal is practically nil, it is no wonder that women became objects of barter in this way. The old men of the tribe have only the young men of their own class to compete with, and since the young men have very little, if any, say in the matter they usually get the women that their elders do not want. One observer[3] recognized this result although he failed to see it as the guiding principle of the system when he said "Reverence for age and authority has greatly aided the elderly men in monopolizing the wives of the class with which they intermarry." Yet having stated this fact he seems to lose sight of its significance and goes on to explain exogamy as having sprung directly from marriage by capture, overlooking the probability that the Australian system followed totemic divisions, indicating that the social organization was already far in advance of the stage of society in which this primitive system of obtaining wives was universal. If food restrictions had been devised and operated for the old men's advantage why should not marriage restrictions operate in the same way—which as we have seen, they do!

Another very general feature which has puzzled observers is the system by which when two people intermarry in their appropriate classes, the children do not belong to the class of either parent but commonly to the alternative class in the phratry or larger division in which the mother belongs. Suppose phratry A is composed of blacksnake and emu classes and

[3] Mathew: "Eaglehawk and Crow," p. 135.

phratry B of kangaroo and black duck. Then an individual in blacksnake class (phratry A) must marry a kangaroo woman of phratry B. Their son would then be black duck. He in turn must marry an emu woman of the opposite phratry (A). His children will be of black-snake class—the same as their grandfather. In this way the children alternate between the classes and in successive generations return to the class of their grandparents.

Apparently no student has seen the reason for this most ingenious system. It makes for tribal unity because it distributes the blood of the tribe among both phratries and all the classes. Any biological advantages such as increased fertility or possession of bodily characters subject to inheritance will be shared equally throughout the tribe. The numbers of individuals too in each class will be numerically well balanced and no class—which means the old men of that class—needs to run out of eligible women as wives.

The phratry system breaks the tribe into exogamous divisions for the benefit of the old men and the crediting of children to alternative classes ensures that tribal unity does not suffer by the division. This effect may be illustrated by imagining the results of an exogamous system without this provision. Supposing we should take a tribe with a fourfold division, A, B, C and D. If A married B and C married D and the children of each union went into the class of either father or mother, then it is obvious that in time there would be two distinct moieties of the tribe only very remotely connected by blood, viz.—the AB group and the CD. If one strain or the other should be stronger then we could readily conceive of one moiety being outnumbered and swamped by the other division. With no intermarriage or blood relationship the tribe might just as well split into two. How different the state of affairs is when A marries B, and their children belong to group D. These in turn marry with group C and their children belong in class A. Here we see in operation a most ingenious plan for preserving and distributing blood relationship throughout the whole tribe. In some tribes without exogamous classes the larger tribal re-

lationships are preserved by a rule that the young man of one locality marries a woman from another. The crediting of offspring in alternating generations to each class is therefore seen to be the counterpart of the phratry system. As with the totemic food restrictions the aborigine kills two birds with the one stone. He establishes a system of food taboos for his own purposes and then uses the system as a means of maintaining tribal unity and coöperation. Similarly he arranges a scheme which ensures a plentiful supply of wives for himself and then incorporates in it a feature which makes the system, in another way, of tribal advantage. Incidentally by the practice of exogamy incest is prevented but it is hard to believe that any considerations of sexual morality were at the back of the plan. Population also may be somewhat restricted by these hindrances to free intermarriage but that is by no means a disadvantage.

Here surely is no evidence of mental obtuseness or primitiveness. The main objection to our theory will probably come from those who are unwilling to concede that the aborigine possesses enough foresight and ingenuity to work out such a system. It seems very probable that much of the Australian's social organization was the result of deliberate planning and design but this does not mean that we must credit him with extreme far-sightedness. Seeing that the wives of the old man frequently belong to three generations, the youngest being just past puberty and probably betrothed to him at birth—it would take but little observation on his part to see the advantages of a system which kept up the numerical balance of females in the intermarrying classes. As regards the genesis of the idea we must remember that women were already objects of barter and the class distribution of offspring as we have outlined it was nothing more than a system of intra-tribal, inter-class barter or exchange of children. Whether the plan was deliberate or not, it works out that way. It should be noted that the writer is not alone in crediting the aborigine with astuteness. Baldwin Spencer, than whom there has been no closer observer and student of native life in Australia, says that he feels "more than ever convinced that matters such as the division of the tribe

into intermarrying groups could well be the result of the deliberate thinking out of a scheme on the part of certain members of the tribe more highly gifted than the common run."

Next we may consider the initiation practices which consist commonly of circumcision and sub-incision or slitting of the male urethra. Since full membership and inclusion in the tribal council is of such tremendous importance in the life of the black it is only natural that there should be initiatory marks and symbols. Since the women are forever excluded[4] from anything remotely resembling authority or high tribal status, the initiatory mark must be distinctively male. In fact the three criteria which must be satisfied are, first, that the mark should be permanent and easily recognized; secondly, that it should be severe as regards its infliction and thus be the badge of fortitude and manliness; and thirdly, that it should be something the women could not copy. Circumcision in part met all three requirements, sub-incision more completely so. The knocking out of teeth, the making of incisions so as to raise huge cicatrices in the fleshy parts of thigh or arm, chest or abdomen, are practised but are not distinctively masculine; women may also endure them and such scars are honourable no matter what the sex of the bearer. But these minor inflictions are naturally of lesser significance than the ones chosen. The idea that sub-incision was devised as a means of preventing conception is incorrect since the operation not only fails of that object but is unnecessary in people who do not hesitate to commit infanticide as a means of regulating population.

But in thus discussing the choice of the initiatory marks we must not neglect their main advantage, which is a psychological one. The effect of the initiation ceremonies which are always carried out with extraordinary mystery, extend over a long period of time and involve the participation of every full member of the tribe, is to impress the young man with the dignity and the power of the old men of the tribal council, who

[4] Except in one or two tribes of the extreme north and southeast of the continent.

have the power to dictate such severe inflictions. Psychologically speaking, nothing is neglected that would help to make such an impression indelible in the mind of the novice so that unquestioning obedience to the tribal elders would forever after be his habit. No masonic rite, no matter how dignified, could be half so effective as these savage ceremonies. In certain tribes the rites begin when the boy is eight or nine and it is the fashion for the father to put up a fight against the emissaries of the tribal council who come to take the boy away. The unfailing effect is, of course, to impress the boy with the idea that even his father's protection is futile against the will of the elders. This impressiveness has also been pointed out by Spencer and Gillen who state that the purpose of secluding the initiate for several days prior to the ceremony is "to impress him strongly with a feeling of the deep importance of strict compliance with the tribal rules and further still with a strong sense of the superiority of the older men who know and are familiar with all the mysterious rites."

Finally we come to one widespread provision which in its effects is as astute as anything to be found in any social organization. This is the universal custom of mother-in-law avoidance. Whenever a man is married or betrothed to a girl, which latter often takes place immediately after her birth, the husband-to-be and the prospective mother-in-law must rigidly avoid each other. There can be no communication between them except through a third party. Under the conditions of daily life the women and children consort together while the males for much of the time are apart, engaged in hunting and carrying out ceremonies. A daughter is therefore very directly under the influence of her mother. Betrothal or marriage is the sign that this dominance has passed out of the hands of the mother to the husband. Thereafter the rule of avoidance ensures that all communication from mother to son-in-law must cease. Hence these two influences never clash since one is ruled out entirely. The husband can use or misuse his wife as he likes without interference from her natural protector. A more direct plan would have been to prohibit mother and

daughter intercourse but since this is impossible the next best plan is adopted.

Enough has been said to show that the most distinctive features of the Australian social organization have behind them an ingenuity of device and purpose which is surprising. A great deal of misunderstanding of their real purport is due to the fact that the student is so thoroughly convinced of aboriginal mental incapacity that he does not look deep enough for a substratum of common sense motivation. There is much in native life which savours strongly of a childish psychology but we must not be deluded into thinking that the aborigine is in all respects as a child or that he is blind to his own advantage. It is not so long since our own fore-parents cleared their skirts of gross superstition but even when they were under the dominance of many irrational beliefs this does not justify us in regarding them as fools. When we consider the nakedness of the Australian's physical environment and the shifts and expedients that mere survival entails it is no wonder that in the social organization of this people there must be unique and effective adaptations which merit appreciation rather than contempt. If we were thrown empty-handed into this environment it would be inevitable that we should ourselves in some particulars imitate some of these aboriginal expedients.

If on the other hand, Australia had been a country which could have supported a large indigenous population then there would have been close tribal contact resulting in rivalry and warfare. Under such circumstances the tribes which divided authority among the elders or which had a marriage system which restricted the number of fighting males would soon be exterminated. Similarly individual prowess and leadership would have high value, and wealth as well as authority would fall into the hands of the tribal defender or champion. An hereditary system would spring up for the simple reason that there would be something to inherit. How could we expect such a system to develop in Australia where conditions are so utterly difficult? Why do we hesitate to give credit to the aboriginal for developing as far as he can under the circum-

stances and to recognize that this development is highly specialized both in degree and direction?

In those parts of Australia where sustenance is easiest—that is to say in the extreme north and in the south-east—we find as we should expect, a changed social organization. It is on these opposite sides of the continent where the class system of intermarriages has either been relaxed considerably or has broken down entirely, the only restriction in the south-east being that young men obtain their wives from a different locality. The initiatory operations are much less severe, being restricted to the knocking out of the front teeth, a practice which is prevalent in the north and central regions but which has lost any special significance. Similarly with the disappearance of the exogamous class system, the practices of circumcision and sub-incision are discontinued, and, since full tribal status is not so significant we find that in some northern tribes women and even children are admitted to the ceremonies. The nearest approach to a system of hereditary chieftainship is also found in some of the tribes of the south-east. Finally in the more favoured parts the totemic system, in common with the rest of the social organization, tended to disappear. All of this seems to prove that even though the origin of these customs may have been grounded in magic and superstition, they did not survive once the advantages associated with them and the necessity for them no longer existed. Further, the close relation between the physical environment and the social organization is demonstrated and the imputed lack of progress in these people is seen to be due, not to any inherent incapacity on their part, but to the fact that their customs have been developed in response to the demands of that environment and are in harmony with it.

If this view is correct then it is futile to examine the social life of the Australian aborigine with a view to observing in it the primitive processes through which civilization has universally passed. Even though the Australian uses the same tools as our Neolithic ancestors it is not necessary to believe that he has similar social behaviour and that therefore in him we are observing the habits that were characteristic of the childhood of

our own race. For example, even though there may be something in Australian aboriginal customs which approaches group marriage we cannot accept this as proof that this was a once universal condition. Herein there is an interesting parallel with physical evolution. Just as the structure of the anthropoid ape represents a highly specialized development and not in any sense a primitive stage in man's evolution, so too the social organization of the Australian aboriginal has pursued its own evolutionary course and may in its maturity represent a distinct divergence from the cultural stem from which our civilization has sprung.

This theory will of course run counter to the popular psychological conception of aboriginal incapacity. There has been, however, very little effort expended by scientists in attempting to discover the mental status of the aborigine by psychological tests. Under the auspices of the Australian National Research Council the writer this year is to undertake a mental survey of some of the tribes in the extreme north-west and centre of Australia. This survey, it is hoped, will throw some light on both the mental capacity and variability of these people.

There is ample evidence that in anything which concerns his own affairs, the aborigine exhibits remarkable powers of observation and memory. The details of ceremonial observances, the songs that are used with various dances, the legends that are imparted to him, if seen or heard a single time are faithfully remembered, while his wonderful powers of visual observation in tracking and food-getting are beyond dispute. When we consider the extreme inadaptability of the mentally defective of our own race we realize that the aboriginal could not possibly survive if he were inherently dull or stupid. His wisdom is not the white man's wisdom and quite possibly is not susceptible to evaluation by the white man's tests and measurements but there is no doubt that he possesses a fund of knowledge and experience that could only be obtained by the exercise of keenness, alertness and reasoning ability—aptitudes indeed which by no means distinguish the mentally inferior.

In conclusion it should be mentioned that while the foregoing psychological interpretation of totemism, exogamy, initiation, etc., represents a new approach to the problem, it rests on a solid basis of recorded observations by competent observers.[5] These observers however, do not appear in any instance to have attempted to relate both the food restrictions and the marriage class system to the one basic condition of old man dominance and government, nor what is perhaps as important, did they recognize the totemic divisions and the scheme of child exchange between classes as the counterpart or complement of exogamy and the means by which tribal unity and numerical balance were preserved.

[5] Space does not admit of the citation of these authorities but these are contained in a longer paper entitled "Psychological Interpretations of Australian Aboriginal Customs."

# THE PORTEUS MAZE TEST AS A DISCRIMINATIVE MEASURE IN DELINQUENCY

LOUISE E. POULL[1] AND RUTH P. MONTGOMERY

*New York City Children's Hospital, Randall's Island*

It is generally agreed that crime and delinquency can be analyzed into many component factors or elements and that the detection of these factors is essential for the reconstruction of criminals and delinquents as well as for the prevention of crime and delinquency. Knowing these elements the clinical psychologist is enabled to detect them before they have created havoc in a personality and to help find the necessary understanding and support to bring about a normal personal integration.

It is obvious to the examiner who interviews delinquent subjects that many of them are puzzled by their own status and are willing to coöperate in plans for readjustment. They have, however, no specific suggestions to make, and looking backward, their explanations are vague. Circumstances have apparently been overpowering to their morale. They appear to be lacking in practical planning ability. It is important, therefore, to measure the development of this function in relation to the general intelligence of an individual in order to anticipate social difficulties.

The Porteus Maze Test is a measure of plannng ability in a restricted situation and its author has claimed that its scores are lower in cases of delinquency. The present study was made to test the validity of this assertion, and is supplementary to the author's own experiments.[2]

The data furnished by the records of the New York City

[1] Acknowledgment is due to Prof. H. E. Garrett for consultation on statistical method.

[2] Porteus, S. D., Studies in Mental Deviation, 1923, Training School Bulletin, Vineland, N. J., pp. 75–116.

Children's Hospital, Randall's Island, enable us to compare two well defined groups: conduct disorder cases and socially adjusted cases. The criteria for conduct disorder are stealing, obstreperous behavior, truancy, fighting, and inability to get along with companions or guardians. Seventy-two cases were found for this group, 30 girls and 42 boys. The control group, numbering 81 cases, 34 girls and 47 boys, contains the law abiding, self controlled, diligent, and coöperative individuals. Doubtful cases have been excluded from the study as well as those in which the physical condition or epilepsy may have affected the test scores.

The chronological age limit for both groups is necessarily 15 years 11 months since social maladjustment is almost the sole reason for admission of subjects above the inbecile grade after age 16 years.

The following tables show the distributions of both groups for chronological age, MA on Stanford-Binet, IQ on Stanford-Binet, MA on Porteus Maze Test, IQ on Porteus Maze Test.[3]

*Chronological ages on Stanford-Binet*

| *Socially adjusted cases* | | | | *Socially mal-adjusted cases* | | | |
|---|---|---|---|---|---|---|---|
| 8–5 | 1 | 12–6 | 6 | 10–6 | 3 | 14–0 | 14 |
| 9–0 | 1 | 13–0 | 9 | 11–0 | 2 | 14–6 | 7 |
| 9–6 | 0 | 13–6 | 11 | 11–6 | 1 | 15–0 | 8 |
| 10–0 | 5 | 14–0 | 7 | 12–0 | 5 | 15–6 | 14 |
| 10–6 | 0 | 14–6 | 9 | 12–6 | 6 | | 72 |
| 11–0 | 5 | 15–0 | 14 | 13–0 | 8 | | |
| 11–6 | 2 | 15–6 | 7 | 13–6 | 4 | | |
| 12–0 | 4 | | 81 | | | | |
| Av. | 13 yrs. 7.15 mos. | | | Av. | 14 yrs. 0.4 mos. | | |
| SD | 20.28 mos. | | | SD | 16.8 mos | | |
| SD av. | 2.253 mos. | | | SD av. | 1.981 mos. | | |

SD diff. 3.00

$$\frac{D}{\text{SD diff.}} \quad 1.75 \text{ SD}$$

1.75 9599 chances in 10,000 that difference is greater than 0

[3] For detail of this investigation see "The Correlation of the Porteus Maze Test and the Healy 11 Pictorial Completion Test with Social Conduct, by Ruth P. Montgomery. M.A. thesis. Columbia University, 1927.

*IQ's on Stanford-Binet*

| *Socially adjusted* | | | | *Socially mal-adjusted* | | | |
|---|---|---|---|---|---|---|---|
| 50 | 7 | 66 | 5 | 42 | 2 | 62 | 9 |
| 52 | 3 | 68 | 4 | 44 | 0 | 64 | 8 |
| 54 | 9 | 70 | 4 | 46 | 0 | 66 | 6 |
| 56 | 5 | 72 | 4 | 48 | 0 | 68 | 1 |
| 58 | 9 | 76 | 6 | 50 | 4 | 70 | 9 |
| 60 | 4 | 78 | 1 | 52 | 1 | 72 | 4 |
| 62 | 10 | | 81 | 54 | 7 | 74 | 2 |
| 64 | 7 | | | 56 | 2 | 76 | 2 |
| | | | | 58 | 8 | 78 | 0 |
| | | | | 60 | 6 | 80 | 1 |
| | | | | | | | 72 |
| Av | | | 63 | Av | | | 63.25 |
| SD | | | 7.73 | SD | | | 7.72 |
| SD av | | | 0.859 | SD av | | | 0.910 |

SD diff. 1.25

$$\frac{D}{\text{SD diff.}} \quad 0.20$$

0.20 5,793 chances in 10,000 of a difference greater than 0

*Mental ages on Stanford-Binet Test*

| *Socially adjusted cases* | | | | *Socially mal-adjusted cases* | | | |
|---|---|---|---|---|---|---|---|
| 6–3 | 2 | 9–3 | 2 | 6–3 | 1 | 8–9 | 9 |
| 6–6 | 5 | 9–0 | 2 | 6–6 | 1 | 9–0 | 2 |
| 6–0 | 3 | 9–9 | 2 | 6–9 | 1 | 9–3 | 7 |
| 7–0 | 1 | 10–0 | 1 | 7–0 | 4 | 9–6 | 2 |
| 7–3 | 7 | 10–3 | 1 | 7–3 | 4 | 9–9 | 3 |
| 7–6 | 8 | 10–6 | 3 | 7–6 | 5 | 10–0 | 3 |
| 7–9 | 6 | 10–9 | 1 | 7–9 | 4 | 10–3 | 2 |
| 8–0 | 9 | 11–0 | 0 | 8–0 | 10 | 10–6 | 5 |
| 8–3 | 6 | 11–3 | 2 | 8–3 | 2 | 10–9 | 1 |
| 8–6 | 9 | 11–1 | 1 | 8–6 | 5 | 11–0 | 1 |
| 8–9 | 6 | 11–9 | 1 | | | | 72 |
| 9–0 | 3 | | 81 | | | | |
| Av | | | 8 yrs. 5.09 mos. | Av | | | 8 yrs. 8.58 mos. |
| SD | | | 15.06 mos. | SD | | | 13.59 mos. |
| SD av | | | 1.673 mos. | SD av | | | 1.602 mos. |

SD diff. 2.32 mos.

$$\frac{D}{\text{SD diff.}} \quad 1.504 \text{ SD}$$

1.504 9,332 chances in 10,000 that difference will be greater than 0

*Mental ages on Porteus Maze Test*

| *Socially adjusted* | | | | *Socially mal-adjusted* | | | |
|---|---|---|---|---|---|---|---|
| 4-0 | 1 | 9-0 | 8 | 4-6 | 1 | 8-6 | 4 |
| 4-6 | 0 | 9-6 | 12 | 5-0 | 9 | 9-0 | 8 |
| 5-0 | 3 | 10-0 | 3 | 5-6 | 5 | 9-6 | 4 |
| 5-6 | 3 | 10-6 | 6 | 6-0 | 7 | 10-0 | 3 |
| 6-0 | 1 | 11-0 | 0 | 6-6 | 8 | 10-6 | 2 |
| 6-6 | 6 | 11-6 | 4 | 7-0 | 5 | 11-0 | 4 |
| 7-0 | 9 | 12-0 | 0 | 7-6 | 5 | 11-6 | 1 |
| 7-6 | 6 | 12-6 | 2 | 8-0 | 6 | | 72 |
| 8-0 | 9 | 13-0 | 1 | | | | |
| 8-6 | 7 | | 81 | | | | |
| Av. | | 8 yrs. 8.16 mos. | | Av. | | 7 yrs. 9.33 mos. | |
| SD. | | 22.14 mos. | | SD. | | 22.44 mos. | |
| SD av. | | 2.460 mos. | | SD av. | | 2.646 mos. | |

SD diff. 3.61

$$\frac{D}{\text{SD diff.}} \quad 3.00$$

3.00 9,986.5 chances in 10,000 that difference will be greater than 0

*The IQ's for the Porteus Maze Test*

| *Socially adjusted cases* | | | | *Socially mal-adjusted cases* | | | |
|---|---|---|---|---|---|---|---|
| 35 | 4 | 80 | 4 | 30 | 4 | 70 | 3 |
| 40 | 2 | 85 | 0 | 35 | 5 | 75 | 4 |
| 45 | 9 | 90 | 2 | 40 | 12 | 80 | 2 |
| 50 | 10 | 95 | 1 | 45 | 6 | 85 | 2 |
| 55 | 6 | 100 | 1 | 50 | 13 | 90 | 1 |
| 60 | 13 | 105 | 0 | 55 | 7 | | 72 |
| 65 | 11 | 110 | 1 | 60 | 8 | | |
| 70 | 8 | | 81 | 65 | 5 | | |
| 75 | 9 | | | | | | |
| Av. | | 64.1 | | Av. | | 55.48 | |
| SD. | | 14.95 | | SD. | | 14.45 | |
| SD av. | | 1.661 | | SD av. | | 1.704 | |

SD diff. 2.38

$$\frac{D}{\text{SD diff.}} \quad 3.62$$

3.62 9,998.4 chances in 10,000 of difference greater than 0

Following are the tables comparing the findings on the Stanford-Binet with those of the Porteus Maze Test.

*Mental ages*

| *Adjusted cases* | | *Mal-adjusted cases* | |
|---|---|---|---|
| Stanford-Binet | | Stanford-Binet | |
| Av. 8 | 5.05 | Av. 8 | 8.58 |
| SD av | 1.673 | SD av | 1.602 |
| Porteus | | Porteus | |
| Av. 8 | 8.16 | Av. 7 | 9.33 |
| SD av | 2.46 | SD av | 2.646 |
| SD diff | 2.975 | SD diff | 3.09 |
| $\frac{D.}{SD\ diff.}$ | 1.03 | $\frac{D}{SD\ diff.}$ | 3.64 |
| 1.03 8,485 chances in 10,000 that difference is greater than 0 | | 3.64 9,998.6 chances in 10,000 that difference is greater than 0 | |

*IQ's*

| *Adjusted* | | *Mal-adjusted* | |
|---|---|---|---|
| Stanford-Binet | | Stanford-Binet | |
| Av | 63 | Av | 63.25 |
| SD av | 0.8589 | SD av | 0.9104 |
| Porteus | | Porteus | |
| Av | 64.1 | Av | 55.48 |
| SD av | 1.761 | SD av | 1.704 |
| SD diff | 1.87 | SD diff | 1.93 |
| $\frac{D}{SD\ diff.}$ | 0.588 | $\frac{D}{SD\ diff.}$ | 4.03 |
| 0.588 7,224 chances in 10,000 that difference is greater than 0 | | 4.03 9,999.68 chances in 10,000 that difference is greater than 0 | |

Inspection of these tables reveals the fact that there is a tendency for the socially adjusted children to make slightly higher scores on the Porteus Maze Test than they make on Stanford-Binet, and conversely, for mal-adjusted children to make lower scores on the Porteus Maze Test than they make on the Stanford-Binet. The mal-adjusted group is slightly older and less variable in mental age and intelligence quotient on both scales. Since there is no correlation between chronological age and mental age on the Porteus Maze scale,[4] we may dismiss this factor from our consideration.

[4] Karpeles, "An Alternate for Stanford Binet." M.A. thesis Columbia University, 1927.

The difference between the findings of these two scales is further emphasized by a study of both groups on the Healy 11 Pictorial Completion Test. The following table shows the comparison of MA and IQ by the Stanford-Binet and Healy 11 Pictorial Completion Test.

*Mental ages*

| *Adjusted cases* | | *Mal-adjusted cases* | |
|---|---|---|---|
| Stanford-Binet | | Stanford-Binet | |
| Av. 8 | − 8.509 | Av. 8 | − 8.58 |
| SD av | 1.673 | SD av | 1.602 |
| Healy 11 | | Healy 11 | |
| Av. 9 | − 6.49 | Av. 9 | − 6.75 |
| SD av | 3.633 | SD av | 4.677 |
| SD diff | 3.999 or 4 | SD diff | 4.94 |
| $\frac{D}{SD\ diff.}$ | 3.35 | $\frac{D}{SD\ diff.}$ | 2.06 |
| 3.35 9,995 chances in 10,000 that difference is greater than 0 | | 2.06 9,803 chances in 10,000 that difference is greater than 0 | |

*IQ's*

| *Socially adjusted cases* | | *Socially mal-adjusted cases* | |
|---|---|---|---|
| Stanford-Binet | | Stanford-Binet | |
| Av. | − 63 | Av. | − 63.25 |
| SD av | 0.8589 | SD av | 0.9104 |
| Healy 11 | | Healy 11 | |
| Av. | − 71.52 | Av. | − 69.37 |
| SD av | 2.17 | SD av | 2.683 |
| SD diff | 2.33 | SD diff | 2.83 |
| $\frac{D}{SD\ diff.}$ | 3.66 | $\frac{D}{SD\ diff.}$ | 2.16 |
| 3.66 9,998.92 chances in 10,000 that the difference is greater than 0 | | 2.16 9,846 chances in 10,000 that the difference is greater than 0 | |

Here we find no difference between the two groups. If we may assume that the Pictorial Completion tests apperception in the social field, we may conclude that the social difficulty of these children is not caused by inability to comprehend or interpret the world about them, but is due, at least in part, to

inability to meet the exigencies of new situations. The anti-social behavior with its concomitant emotional disturbances and defense reactions, may be a result of panic rather than malicious intent.

In the case of children who steal, numbering 17 in this study, the median mental age is 9 years 0 months in both Stanford-Binet and Porteus Maze tests. The conclusions we have drawn are therefore not applicable to them. The number is, however, too small for final generalization.

Illustrations are not wanting in the life of the institution to exemplify the relative social values of ratings on verbal intelligence tests and ratings on the Porteus Maze Scale. A girl whose composite rating on all tests taken places her in the high grade group is given an opportunity to qualify as waitress in the nurses dining room. Her behavior in the schoolroom and on the ward has been satisfactory. After a short period she is referred to the laboratory because a decided change in behavior has been observed. She now is described as excitable and pugnacious, incompetent and demoralizing to the service in which she is placed. Examination of her records shows that although her logical thinking is at 10 years her planning power, shown in maze tests, is at 7 years. Another girl whose verbal intelligence is not over 8 years, but who rates 11 years on mazes, and who was a difficult behavior problem in the academic school, proves to be a good waitress and is now reported as successful in extra institutional work. Instances of this type could be multiplied.

Girls and boys of age 16 years or over who are admitted to the institution in consequence of incompetency or social difficulties have, as a group, markedly low scores on the maze tests.

A study of various types of delinquency and also of upper levels of intelligence in regard to these differences in test scores would further clarify this discussion.

# THE PREDICTIVE VALUE OF MENTAL TESTS THAT SATISFY SPEARMAN'S TETRAD CRITERION

E. J. ASHER

*University of Kentucky*[1]

The object of this paper is to give the results obtained through the application of Spearman's two factor theory and tetrad criterion to the practical problem of constructing an "intelligence" test for the prognostication of success in college work. A five-year testing program at the University of Texas made possible, at the same time, a study of the relative prognostic value of tests in which the tetrad criterion was satisfied and those where the criterion was not satisfied.

This five-year testing program was initiated at the University of Texas in the fall of 1923. Under this program the freshmen of the College of Arts and Sciences were to be given an intelligence examination each year beginning with the 1923–24 class and continuing for a period of five years. A committee, of which Dr. F. A. C. Perrin of the department of psychology was chairman, was appointed to construct a test and to administer it to the 1923–24 freshman class. The test constructed and administered by this committee was known as the University of Texas Freshman Psychological Examination.[2]

At its initiation, this testing program had as its sole aim the prognostication of success in college work. "Our interest lies solely in the factors, whatever they may be, that make for success in college work. . . . . So far as this paper is concerned, the test is valuable to the extent that it predicts scholastic achievement."[3] As a necessary corollary the test became

[1] This work was done while the writer was at the University of Texas.

[2] University of Texas Bulletin. No. 2445, December 1, 1924.

[3] Ibid.

an experiment, an attempt to determine the relative predictive value of various kinds of test items, and the nature of the abilities tested by these items. This aim was dominant throughout the last three years of the testing.

The test constructed and administered to the freshman class of 1923–24 contained such items as synonyms, opposites, analogies, items found in the ordinary intelligence test, and in addition it contained items pertaining to school subjects, such as English. The tests constructed and administered to the freshman classes in 1924–25 and 1925–26 placed more emphasis upon the items pertaining to school information. These tests included such items as History, Science, English, and Syllogisms. The 1924 and 1925 tests consisted of 100 problems of 10 different kinds arranged in omnibus fashion so that the problems of every tenth place were similar. The time limit for each test was 50 minutes.

In spite of the fact that these tests were loaded with items pertaining to school information, the coefficients of correlation between the scores on the tests and average grades during the first term in college were in the neighborhood of the average coefficients obtained with other intelligence tests. The coefficients of correlation were 0.49, 0.447, and 0.438 for the 1923, 1924, and 1925 tests respectively.

In an effort to determine the relative predictive value of the kinds of test items used and the nature of the abilities tested by these items, the 1925 test was subjected to a thorough statistical analysis by the present writer, who at that time was placed in charge of the test. This analysis consisted in finding intercorrelations between the test items, correlations between each test item and average scholarship, correlations between each test item and each school subject. These correlations revealed that the items pertaining to school information predicted success in their respective school subjects with a comparatively high degree of accuracy. The History test predicted history grades very well. The Science test predicted science grades very well. But these tests did not predict grades in other subjects, nor did they predict total scholarship as well

as such tests as synonyms, opposites, and reading comprehension. The application of Spearman's tetrad equation to the intercorrelations of the test items revealed much overlapping in the abilities tested by the items. The tetrad equation did not hold throughout the table of correlations.

With these facts at hand, a new test was constructed in 1926. This test consisted of 100 problems of five different kinds arranged in ascending order of difficulty in omnibus fashion. The time limit was reduced from 50 minutes to 45 minutes. The different kinds of problems can be referred to as opposites, synonyms, general information, reading comprehension, and mathematics. All problems were put into the multiple choice type of answer with the exception of the mathematics problems. Each possible answer was numbered and the subjects were instructed to place the number of the correct answer in the parenthesis at the right. In this way all the answers appeared in one row at the right hand edge of the page. A sample of each kind of problem with the instructions is shown below.

*The opposite of* exhume *is* (*write the number of the correct word in the parentheses at the right*)

(1) curse (2) swear (3) bury (4) exhale. ( )

*Use the margin for figuring and write the answer in the space indicated:*
What part of 8 equals one-half of 12? (*Answer:*———)

*Place in the parentheses at the right the number of the nearest correct statement:*

An oratorio is (1) The name of a kind of musical composition (2) The name of a bird (3) A kind of speech (4) The name of one of Shakespeare's characters. ( )

*The synonym, or word corresponding in meaning, for* expel *is* (*write the number of the correct word in the parentheses at the right.*

(1) inject (2) eject (3) shoot (4) admit. ( )

*Write in the parentheses at the right the number of one of the four numbered statements that comes nearest expressing the thought of the quotation.* ( )

"The person who writes his autobiography is apt to distort the facts of his life."

1. A man's life is distorted when he writes his autobiography.
2. One who writes his autobiography always distorts the facts.

3. A person who writes a history of his life is liable to distort the facts somewhat.
4. A man distorts the facts of his own life.

This test was administered to 805 Arts and Sciences freshmen at the beginning of the fall term of 1926. The results were then subjected to a statistical analysis. The corrected reliability coefficient, between the odd and even problems of the test, was 0.90, p.e. 0.005. In order to determine the nature of the abilities tested by the different test items, Spearman's tetrad criterion was applied to the intercorrelations of the five test items. The intercorrelation table is shown below.

*Intercorrelation table*

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1. Opposites | | 0.473 | 0.551 | 0.625 | 0.548 |
| 2. Mathematics | 0.473 | | 0.400 | 0.420 | 0.436 |
| 3. Information | 0.551 | 0.400 | | 0.563 | 0.393 |
| 4. Synonyms | 0.625 | 0.420 | 0.563 | | 0.582 |
| 5. Reading | 0.548 | 0.436 | 0.398 | 0.582 | |

When the tetrad equation, $r_{ap} \times r_{bq} - r_{aq} \times r_{bp} = 0$,[4] was applied to this table of correlations, the following tetrad differences were obtained:

$$r_{13} \times r_{24} - r_{14} \times r_{23} = -0.018$$
$$r_{13} \times r_{25} - r_{15} \times r_{23} = 0.021$$
$$r_{14} \times r_{25} - r_{15} \times r_{24} = 0.042$$
$$r_{12} \times r_{34} - r_{14} \times r_{23} = 0.016$$
$$r_{12} \times r_{35} - r_{15} \times r_{23} = -0.033$$
$$r_{14} \times r_{35} - r_{15} \times r_{34} = -0.063$$
$$r_{12} \times r_{34} - r_{13} \times r_{24} = 0.035$$
$$r_{12} \times r_{45} - r_{15} \times r_{24} = 0.045$$
$$r_{12} \times r_{45} - r_{15} \times r_{34} = 0.012$$
$$r_{12} \times r_{35} - r_{13} \times r_{25} = -0.054$$
$$r_{12} \times r_{45} - r_{14} \times r_{25} = 0.003$$
$$r_{13} \times r_{45} - r_{14} \times r_{35} = 0.075$$
$$r_{24} \times r_{35} - r_{25} \times r_{34} = -0.080$$
$$r_{23} \times r_{45} - r_{25} \times r_{34} = -0.012$$
$$r_{23} \times r_{45} - r_{24} \times r_{35} = 0.067$$

[4] Spearman, Carl: The Abilities of Man. The Macmillan Company, 1927, p. 73.

According to Spearman's tetrad criterion, whenever the tetrad equation holds throughout any table of correlations, every individual measure of every ability can be divided into two independent parts, one a general factor ($g$), and the other a specific factor ($s$). If $g$ and $s$ exist in a table of correlations without overlap in the $s$'s, the tetrad differences should tend to have just such values as would result from sampling errors alone.[5] If the above tetrad differences are just those differences that are to be expected from sampling errors, the median tetrad difference and the probable error of the tetrad differences should be about equal, and the largest tetrad difference should not be more than five times the magnitude of the probable error.[6] The probable error of the above tetrad differences is 0.031.[7] The observed median tetrad difference is 0.035. All of the differences are within plus and minus 3 times the probable error. This indicates that the tetrad equation holds throughout the table of correlations, and that every individual measure of every ability in the table can be divided into two independent parts, $g$ and $s$, $g$ being common to all of the abilities tested.

With this fact in mind, it is significant to note that the coefficient of correlation between the 1926 test and school grades was 0.605. This fairly high degree of relationship between fall term grades and the combined scores of the five test items whose intercorrelations depend upon the general factor seemingly indicates that scholarship is also dependent in large part upon the general factor.

This dependence of scholarship upon the general factor was further indicated by correlating average scholarship for the fall term with each of the test items, including it in the above intercorrelation table as a variable, and calculating the tetrad differences in the new correlation table thus constructed. The largest difference in this new table was 5.2 times the probable error of the tetrad differences, while 30 of the 45 differences

[5] Ibid., p. 140.
[6] Ibid., p. 141.
[7] Ibid., Appendix, p. xi, formula 16A.

were above the probable error. While all of the tetrad differences do not rigidly conform to those that would be expected from sampling errors, the agreement is a striking one when it is remembered that average scholarship is a complex of five different school subjects under as many different instructors. Looking at the matter from another angle, the largest tetrad difference, which was 0.1361, can hardly be considered significant when it is just 5.2 times the probable error. All of the other tetrad differences were within plus and minus 4 times the probable error. It would seem, therefore, that scholarship is dependent in large part upon *g*. The amount of correlation between scholarship and *g* is indicated by an average coefficient of 0.651.[8]

It is also significant, in this connection, to compare the prognostic value of the 1926 test with that for each of the previous tests. The coefficients between average grades and each of the four tests are as follows:

| | r | p.e. |
|---|---|---|
| 1923 | 0.49 | 0.018 |
| 1924 | 0.447 | 0.018 |
| 1925 | 0.438 | 0.019 |
| 1926 | 0.605 | 0.016 |

It will be remembered that the tetrad equation when applied to the intercorrelations of the test items of the 1925 test did not hold. The 1925 test was a duplicate of the 1924 test as far as kinds of test items were concerned, so that it can be concluded that the tetrad equation would not have been satisfied in the 1924 test. Now, since the prognostic value of these two tests is lower by almost 20 points than that for the 1926 test in which the correlations between the test items depend upon *g* alone, it would seem that a higher correlation between test scores and scholarship is to be expected if the correlations between the test items depend upon *g* alone than if they do not.

This finding is further supported by the results of the 1927

[8] Ibid., Appendix, p. xvi.

test. The tetrad equation when applied to the intercorrelations of the five test items of the 1927 test (not the same items used in the 1926 test) was satisfied. The correlation of the 1927 test scores and scholarship for the first semester was .58, much the same as that for the 1926 test.

It would seem, therefore, that best results are to be obtained in predicting school achievement with a test in which the abilities measured obey the two factor theorem.

Thus the initial aim of the testing program at the University of Texas, the prognostication of success in college work, has been considerably furthered by the construction of tests in which the abilities measured are dependent upon *g*. If these results are taken only tentatively, they point very definitely to the fruitfulness of Spearman's two factor theory as a guide in current testing work, and to his tetrad technique as a method of evaluating current tests.

## SUMMARY

A five year testing program was initiated at the University of Texas in the fall of 1923. The test used in this testing work was the University of Texas Freshman Psychological Examination.

The aim of this testing program was to determine the relative predictive value of various kinds of test items, and the nature of the abilities tested by these items.

The coefficients of correlation between the test scores and scholarship were 0.49, 0.447, 0.438, 0.605, and 0.58 for the 1923, 1924, 1925, 1926, and 1927 forms of the test respectively.

Spearman's tetrad criterion was satisfied when applied to the intercorrelations of the 1926 and 1927 forms of the test. It was not satisfied in the 1924 and 1925 forms.

A higher correlation between test scores and scholarship is to be expected, seemingly, if the correlations between the test items depend upon *g* than if they do not.

The correlations of scholarship with the various test items and with total test scores seemingly indicate that school grades are in large part dependent upon *g*.

# SEX DIFFERENCES IN HANDWRITING

M. EUSTACE BROOM, BLANCHE THOMPSON, AND M. THELMA BOUTON

*State Teachers College, San Diego, California*

There are striking differences between the handwriting of different individuals. Tests of performance of course do not reveal why one individual's handwriting differs from that of another. It has seemed that differences may be due in part to sex of the writers. Several investigations have been made to determine the significance of sex in handwriting.[1] This report presents further data bearing on the problem, and data on the problem of whether women are superior to men in estimating sex of writers of given samples of penmanship.

## PROCEDURE

Since the act of writing is an illustration of a developed sensori-motor habit, and since the process has been made automatic by continual practice in previous years, it is possible to obtain samples of penmanship from college students which are typical of the writing of the individual. Little attention is given to differences in handwriting of members of either sex group studied in this report; on the contrary all attention has been centered on sex differences.

Samples of handwriting were obtained from all members of a class in educational psychology taught by the first writer in the State Teachers College of San Diego during the second semester of 1927–1928. In order that the samples might contain uniform content, the sentence, "Now is the time for all good men to come to the aid of the party," was used as "copy." The

[1] Starch, Daniel: "Educational Psychology. Revised Edition." New York, The Macmillan Company, 1927. Pages 348–9.

samples were written on notebook paper, so no hint as to sex was given by stationery used. Forty samples were selected at random from the assortment obtained. Eighteen of these samples were written by men, and twenty-two by women. These samples were posted on a bulletin board with these directions:

Rate each sample as M (masculine) or F (feminine) by number as follows:

Illustration: 1. M
2. F
3. F
Etc.

Note: There are more samples of feminine handwriting than of masculine handwriting, so this judging process need not degenerate to a process of guessing in the selection of the final samples.

Two separate judgments were obtained from twenty-four persons, eight men and sixteen women, two of whom were trained penmanship teachers. These judgments were obtained approximately two weeks apart, allowing for normal forgetting of rating on the first sample. The samples were taken down after the first judging and no comment was made at any time as to the second rating which was given as a "surprise." The numbers on the samples also were changed for the second rating.

## FINDINGS

The average correct guess in terms of per cent of all guesses is shown in table 1, together with certain other measures.

Table 1 suggests a number of interesting conclusions. The mean (average) scores verify the findings of Binet, Starch, and Downey who have found that untrained subjects can determine sex of writers in from 65 to 75 per cent of given specimens.[2] In the present study, no significant differences were observed in the percentages of correct judgments of men and of women. The probable errors of the several means indicate that the chances are even (one to one) that the true average lies within Starch's limits in four cases, and almost within those limits in

[2] Starch, D.: Loc. cit., page 348.

all six instances. (It must be kept in mind that the probable error is the error which may be expected according to the theory of probability. It has nothing to do with the errors due to inadequate or defective data, incorrect computation, or slipshod thinking in regard to the problem under analysis.) While both the men and the total group of judges guessed sex of writers of

TABLE 1

*Percentile measures summarizing the judgments of sex from handwriting*

| SEX OF JUDGES | NUMBER | JUDGMENT OF SERIES | MEAN | P.E.$_M$ | STANDARD DEVIATION | C. OF V. (PEARSON) |
|---|---|---|---|---|---|---|
| Men | 8 | 1 | 70.0 | 5.69 | 23.87 | 34.1 |
| | | 2 | 68.3 | 5.79 | 24.29 | 35.6 |
| Women (other than trained penmen) | 14 | 1 | 69.5 | 4.49 | 24.88 | 35.8 |
| | | 2 | 71.0 | 4.72 | 26.17 | 36.9 |
| Both men and women (including trained penmen) | 24 | 1 | 68.0 | 3.05 | 22.13 | 32.5 |
| | | 2 | 67.7 | 3.14 | 22.79 | 33.7 |

TABLE 2

*Measures of reliability of judgments*

| JUDGMENTS BY | NUMBER | r (PEARSON) | P.E.$_r$ | $k$ |
|---|---|---|---|---|
| Men | 8 | 0.580 | 0.159 | 0.814 |
| Women (other than trained penmen) | 14 | 0.746 | 0.080 | 0.666 |
| All judges (including trained penmen) | 24 | 0.712 | 0.068 | 0.702 |

the forty specimens correctly less often on the second than on the first judging, the variability of judgments was greater on the second than on the first judgment. Both the standard deviation, an absolute measure, and the coefficient of variation, a relative measure, were larger for the second series of judgments for all groups of judges.

Table 2 presents measures summarizing the consistency with which the several groups of judges guessed sex of writers correctly.

It can be seen from a study of table 2 that the group of untrained women judges was more consistent in judging sex of writers correctly than a smaller group of untrained men judges. Indeed, the untrained women judges were more consistent in rating sex of writers of given specimens of penmanship twice than was the total group, including two trained penmanship teachers. The values of the coefficient of alienation ($k$) which measures presence of disagreement as the coefficient of correlation ($r$) measures presence of agreement are such that we know that the group of eight men judges would agree in their ratings in the first and second series of judgments 59.3 times in each 100 instances; the women judges, 66.7 times in each 100 instances; and the total group 64.9 times in each 100 instances. It is obvious, therefore, that judgments of untrained persons are more valuable in determining sex of writers of given specimens of penmanship as to accuracy (and consistent accuracy) of judgment than is the law of chance (50 times in 100).

## ANALYSIS OF CHARACTERISTICS OF SAMPLES "MOST OFTEN JUDGED"

An effort was made in this study to determine whether samples of feminine handwriting most frequently judged feminine differed materially from samples of masculine handwriting most frequently judged feminine; whether samples of masculine handwriting most frequently judged masculine differed materially from samples of feminine handwriting most frequently judged masculine, and so on. Three samples of feminine handwriting most frequently judged feminine were grouped together, as were three samples of feminine handwriting most frequently judged masculine, three samples of masculine handwriting most frequently judged feminine, and three samples of masculine handwriting most frequently judged masculine. These several groups are presented in figure 1.

Several tables showing the judgments of the several groups of judges on those samples "most frequently judged" follow (tables 3 to 6).

A study of samples numbered one to three and ten to twelve

Now is the time for all good men to come to the aid of the party.
Sample 1. Feminine handwriting, G. C.
Now is the time for all good men to come to the aid of the party.
Sample 2. Feminine handwriting, F. A.
Now is the time for all good men to come to the aid of the party.
Sample 3. Feminine handwriting, E. I. F.
Group 1. Samples of feminine handwriting most frequently voted feminine.
Now is the time for all good men to come to the aid of the country.
Sample 4. Feminine handwriting, H. S. G.
Now is the time for all good men to come to the aid of the party
Sample 5. Feminine handwriting, F. B. H.
Now is the time for all good men to come to the aid of the party
Sample 6. Feminine handwriting, C. F. V.
Group 2. Samples of feminine handwriting most frequently voted masculine.
Now is the time for all good men to come to the aid of the party
Sample 7. Masculine handwriting, J. K.
Now is the time for all good men to come to the aid of the party
Sample 8. Masculine handwriting, E. B. C.
Now is the time for all good men to come to the aid of the party.
Sample 9. Masculine handwriting, J. E. S.
Group 3. Samples of masculine handwriting most frequently voted masculine.
Now is the time for all good men to come to the aid of the party
Sample 10. Masculine handwriting, W. K.
Now is the time for all good men to come to the aid of the party.
Sample 11. Masculine handwriting, H. B. B.
Group 4. Samples of masculine handwriting most frequently voted feminine.

FIG. 1. SAMPLES OF PENMANSHIP "MOST OFTEN JUDGED"

inclusive will show that regularity, curves, conventional form, and uniformity of slant are characteristic of the samples judged most frequently to be feminine handwriting. Similarly, a study of samples numbered four to nine will show that irregu-

TABLE 3

*Comparison of ratings on samples of penmanship by men which were most frequently rated masculine*

| JUDGMENTS BY | SAMPLE NO. 7 | | SAMPLE NO. 8 | | SAMPLE NO. 9 | |
|---|---|---|---|---|---|---|
| | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments |
| Men | 6 | 75 | 6 | 75 | 6 | 75 |
| Women (other than trained penmen) | 12 | 85.71 | 13 | 92.86 | 13 | 92.86 |
| All judges (including trained women penmen) | 20 | 83.33 | 21 | 87.50 | 20 | 83.33 |

TABLE 4

*Comparison of ratings on samples of penmanship by men which were most frequently rated feminine*

| JUDGMENTS BY | SAMPLE NO. 10 | | SAMPLE NO. 11 | | SAMPLE NO. 12 | |
|---|---|---|---|---|---|---|
| | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments |
| Men | 5 | 62.5 | 2 | 25 | 1 | 12.5 |
| Women (other than trained penmen) | 5 | 35.71 | 1 | 7.14 | 1 | 7.14 |
| All judges (including trained penmen) | 10 | 41.67 | 5 | 20.83 | 2 | 8.33 |

larity, unconventional form, angles, and individual slant are characteristic of the samples judged most frequently to be masculine handwriting. Observation of the twelve samples will show that those samples of handwriting by men most frequently judged masculine are very different from those sam-

ples of handwriting by women most frequently judged feminine; that the samples of handwriting by men most frequently judged feminine resemble the samples of handwriting by women most frequently judged feminine; that the samples of handwriting by

TABLE 5

*Comparison of ratings on samples of penmanship by women which were most frequently rated feminine*

| JUDGMENTS BY | SAMPLE NO. 1 | | SAMPLE NO. 2 | | SAMPLE NO. 3 | |
|---|---|---|---|---|---|---|
| | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments |
| Men | 7 | 87.5 | 7 | 87.5 | 6 | 75 |
| Women (other than trained penmen) | 14 | 100 | 14 | 100 | 14 | 100 |
| All judges (including trained penmen) | 23 | 95.83 | 22 | 91.67 | 22 | 91.67 |

TABLE 6

*Comparison of ratings on samples of penmanship by women which were most frequently rated masculine*

| JUDGMENTS BY | SAMPLE NO. 4 | | SAMPLE NO. 5 | | SAMPLE NO. 6 | |
|---|---|---|---|---|---|---|
| | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments | Correct judgments | Per cent of all judgments |
| Men | 0 | 0 | 3 | 37.5 | 2 | 25 |
| Women (other than trained penmen) | 3 | 21.43 | 2 | 14.29 | 7 | 50 |
| All judges (including trained penmen) | 5 | 20.83 | 5 | 20.83 | 9 | 37.50 |

women most frequently judged masculine resemble the samples of handwriting by men most frequently judged masculine; and that the samples of handwriting by women most frequently judged masculine resemble the samples of handwriting by men most frequently judged feminine. Such findings seem to

justify the following conclusions: feminine characteristics were found in the handwriting of men which caused it to be judged feminine, and masculine characteristics were found in the handwriting of women which caused it to be judged masculine. Apparently, then, there are certain characteristics which differentiate feminine handwriting from masculine handwriting.

## UTILITY OF FINDINGS FOR TEACHING PURPOSES

It is apparent from the results and conclusions based upon those results that there are sex differences in handwriting, and that it is possible to determine sex of writers correctly roughly two times in every three judgments when the sex of the writer is unknown. It has been shown also that such judgments can be relied upon to be reasonably consistent, so it is apparent that the differences are real, i.e., such that untrained judges will identify as being characteristically masculine or feminine on consecutive judgments. It has not been shown that such differences result from lack of training, or that training causes inverted judgments, that is, samples of handwriting of men to be judged feminine or samples of handwriting of women to be judged masculine. Further studies involving the judging of handwriting by both trained and untrained persons of both sexes might yield illuminating evidence on this point. The present study is suggestive, but it does not yield findings of value in determining teaching practices.

# AN INVESTIGATION OF SOME SIMPLE SPEED ACTIVITIES

KARL C. GARRISON

*North Carolina State College of Agriculture and Engineering*

Recent developments and uses of group tests have revealed the need for further investigations of the factors contributing to success in the performance of the various tests. A brief survey of the group tests of intelligence and the various prognostic tests reveals the fact that among the procedure for giving such tests one usually finds a definite time limit within which the individual or individuals tested must complete the task. It is here assumed that the rate of work is a factor to be considered in measuring the particular ability. There has been a tendency more and more to simplify testing and this has resulted in keeping uniform the rate of response.

Highsmith[1] summarizes the more important investigations bearing on the general problem of the relation between rate of response and intelligence. He concludes from a summary of the investigations until 1924 as follows:

1. There is a comparatively low correlation between rate of performing simple tasks and intelligence. The correlations range from 0.00 (Brown) to 0.61 (Burt). Most of these fall around 0.30 to 0.40 (Wyatt, Brown, McCall).
2. There is a high correlation between rates of performing different simple tasks. Brown gets 0.80 to 0.82 and Gates 0.95.
3. Individuals show considerable differences in rate of response to words presented (Anderson).
4. Group tests and Binet-Simon tests give correlations averaging about 0.55 to 0.65 (Root, Gates, Garrison).
5. In all cases linguistic tests give higher correlations with intelligence than do non-linguistic (Gates).

---

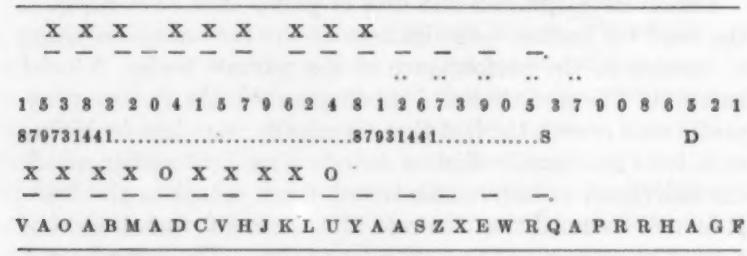
[1] Highsmith, J. A., The Relation of the Rate of Response to Intelligence, *Psychol. Monog.*, 1924, 34, No. 3.

6. The Army experiment and the Ruch and Koerth study do not give sufficient evidence of the closeness of the relation assumed in timed tests between rate of response and intelligence.

Highsmith concludes from his study that the rate of response was not a good measure of intelligence.[2] Also simple linguistic rate tests were found to be as good a measure of intelligence as the National Intelligence test.

TABLE 1

*Illustrating the processes involved in the speed tests*


X X X X X X X X X
1 5 3 8 3 2 0 4 1 5 7 9 6 2 4 8 1 2 6 7 3 9 0 5 3 7 9 0 8 6 5 2 1
8 7 9 7 3 1 4 4 1 . . . . . 8 7 9 3 1 4 4 1 . . . . . S D
X X X X O X X X X O
V A O A B M A D C V H J K L U Y A A S Z X E W R Q A P R R H A G F

Various other attacks have been made on the general problems connected with rate of response. Sisk[3] and Bernstein[4] conducted independently two rather thorough investigations of speed. The conclusions from both investigations are that there is no reliable correlation existing between the rate of response in different types of activities.

The writer found from a study made with college subjects,

[2] The criteria used by Highsmith was (*a*) Intelligence—Stanford Binet tests. (*b*) Group mental ability tests—National Intelligence tests. (*c*) Non-linguistic rate tests. Pressey Primary and Kingsbury combined. (*d*) Linguistic rate tests. Pressey Intermediate classification tests grades 3–6, and the Woodworth-Wells Easy Directions tests. See his Relation of the Rate of Response to Intelligence, *Psychol. Monog.*, 1924, 34, No. 3.

[3] Sisk, T. K., The Interrelations of Speed in Simple and Complex Responses, *George Peabody College for Teachers Contributions to Education*, 1925.

[4] Bernstein, E., Quickness and Intelligence, *Brit. J. of Psychol. Monog. Suppl.*, 1924, 31, 1–55.

using materials from the National Intelligence tests and some simple arithmetical processes that correlations were around 0.80 ±0.05 between the ability to do the various activities in a mixed order and in continuous order.[5] In a more recent study[6] he found positive correlations ranging from 0.17 ±0.10 to 0.40 ±0.08 between the speed in four different types of performances and Peterson's[7] Rational Learning Test. These performances are: Cancellation of A's, cancellation of 2's and 7's; completion of XO series; and some different exercises. Table 1 illustrates the different types of activities.

The purpose of this study is to aid in answering such questions related to rate of response, as: To what extent do the correlations between rate of response and intelligence vary with linguistic and non-linguistic materials, with responses in which the motor element varies, and with responses of varying degrees of complexity? To what extent is the degree of identity in the type of activity a factor in enabling one to predict a subject's ability from one task to another? What is the relation between simple speed activities and power tests presenting problems of thinking in which the rational organization of the situation is involved?

The data compiled for this study were gathered from a class of sixty students in general psychology, although the number taking the different tests varied slightly. The data used in studying the ability for rational organization and higher thought activities were taken from results on the Otis Self Administering Tests of Mental Ability, and also from twenty-four analogies involving forms as used by Thurstone in the College Entrance Tests for the American Council on Education. The tests used in which speed was the prominent factor, were: (1) speed of cancellation of fours from a group of printed digits (100 sec-

[5] Garrison, K. C. A Study of Mental Set. *Unpublished Masters Thesis*, University of N. C., 1926.

[6] Garrison, K. C. An Analytic Study of Rational Learning. *George Peabody College for Teachers Contributions to Education*. No. 44, 1928.

[7] Peterson, Jos., Experiments in Rational Learning. *Psychol Rev.*, 1918, 25, 443–467.

onds); (2) speed of cancellation of c's, i's, m's, and r's (100 seconds); (3) speed of substituting letters for numbers (130 seconds); (4) speed of tallying (++++) (two intervals of 95 seconds each); (5) speed of card sorting (40 cards, 10 trials each); (6) speed of card distribution (40 cards, 5 trials each; (7) speed of vocal responses in reading such materials as objects, colors, letters, forms, numbers, and words (6 pages $8\frac{1}{2}$ x $10\frac{1}{2}$) placed in continuous order, i.e., all of one type of reading material placed together; (8) speed of vocal responses for the

ABLE 2

*Correlations between speed and higher mental ability*

| SPEED TESTS | OTIS S. A. TESTS | FORM ANALOGIES |
|---|---|---|
| Reading mixed | 0.28 ±0.08 | 0.34 ±0.08 |
| Reading continued | 0.18 ±0.09 | 0.17 ±0.09 |
| Substitution of letters for numbers | 0.01 ±0.09 | 0.21 ±0.09 |
| Cancellation of 4's | −0.14 ±0.09 | 0.28 ±0.08 |
| Cancellation of c's, i's, m's, and r's | −0.09 ±0.09 | 0.21 ±0.09 |
| Marking tallies | 0.01 ±0.09 | 0.22 ±0.09 |
| Card sorting | −0.02 ±0.09 | −0.10 ±0.09 |
| Card distribution | −0.22 ±0.09 | −0.20 ±0.09 |

reading materials just mentioned except in this case the six pages of printed materials are in mixed order, e.g., there is probably a word followed by a form followed by an object (as a dog or chair) followed by another word, and then by some color, etc.

Results here are taken from the speed tests in terms of the time element. A stop watch was used in recording the time factor in the case of the reading tests. In the case of all other tests the results were recorded in terms of work done per unit of time. Correlations were found between the speed tests and the two tests used as a measure of intelligence. The analogy test is intended as a power test, and ample time was given in this test for all subjects to complete the work given. This test has been given by the writer to several classes in psychology and the

results have been found to give correlations ranging from 0.35 ±0.06 to 0.50 ±0.04 with grade points.[8] Table 2 gives the correlations found between the various speed tests and the two tests involving to a larger extent the higher mental processes. The correlations with the speed tests are consistently low, and only for the speed of vocal responses is there a reliable relationship indicated.

TABLE 3

*Intercorrelations between the various speed activities**

| | READING MIXED | READING CONTINUED | SUBSTITUTION OF LETTERS FOR NUMBERS | CANCELLATION OF C'S, I'S, M'S AND R'S | CANCELLATION OF 4'S | TALLYING | CARD SORTING | CARD DISTRIBUTION |
|---|---|---|---|---|---|---|---|---|
| Reading mixed | | 0.77 | 0.38 | 0.10 | 0.15 | 0.13 | 0.20 | −0.07 |
| Reading continued | | | 0.41 | 0.05 | 0.26 | −0.05 | 0.19 | −0.08 |
| Substitution | | | | 0.47 | 0.70 | 0.19 | 0.30 | 0.28 |
| Cancellation of c's, i's, m's, r's | | | | | 0.26 | 0.21 | 0.16 | 0.18 |
| Cancellation of 4's | | | | | | 0.32 | 0.22 | 0.09 |
| Tallying | | | | | | | 0.51 | 0.24 |
| Card sorting | | | | | | | | 0.20 |

* The number of subjects performing the different speed activities ranged from 50 to 58.

Intercorrelations were computed for the various speed tests. These correlations are given in table 3. The speed tests might be classified roughly as follows: The two reading activities are tests of speed of vocal responses (linguistic responses); the card sorting and card distribution activities are speed tests of a sensori-motor type in which the motor element is quite prom-

[8] For a method of arriving at grade points see Peterson, Jos., The Rational Learning Test Applied to Eighty-one College Students, *J. Educ. Psychol.*, 1920, 11, 137–150; Garrison, K. C., An Analytic Study of Rational Learning, *George Peabody College for Teachers' Contributions to Education*, No. 44, 1928.

inent; the tallying, substitution, and cancellation tests are also of a sensori-motor type, but in the case of these the sensori element plays an important rôle. At least in the last named classification the activity is of a pencil performance type, which would in a measure limit the motor element involved. The activities can again be arranged roughly into a hierarchy of performances, ranging from those involving almost wholly language activities to those involving in the main motor activities. This arrangement is the one given in table 3. A closer analysis of these correlations reveals a conclusion in which those activities most similar in nature with respect to the language or motor type are most closely related.

The conclusions to be drawn from this study, or that are indicated from it, may be briefly summarized as follows:

1. Correlations between the speed in a performance and higher mental ability depend upon the type of performance. Those performances in which the motor element is prominent give slight negative correlations; but in no case do we find a high positive correlation.

2. The results from the Form Analogies are more closely related to the speed tests than the results from the Otis test. The difference between these results is not sufficient to draw any valid conclusions. Since the Analogy problems constitute a power test this would indicate that the time factor on the Otis test does not show this test to be measuring the same speed elements as are measured by the speed tests. The Analogy power test is even a better measure than the time factor on the Otis test of these same speed elements of the speed tests.

3. A correlation of 0.48 $\pm$0.07 was found between the Otis test and the Form Analogies.[9]

4. The relations between the various speed tests depend partially upon the identity of the factors with respect to their being of the language or motor type.[10]

[9] This is not referred to in the tables or elsewhere in this study.

[10] A careful analysis of the results presented in table 3 reveals the validity of this conclusion.

# OBJECTIVE MEASURES OF ABILITY AND ACHIEVEMENT IN GENERAL PSYCHOLOGY[1]

HOWARD P. LONGSTAFF AND JAMES P. PORTER

*Ohio University*

In recent years there has been much interest shown in, and work done on methods of measuring educational achievement. At the same time there has been a remarkable amount of time and energy devoted to the devising of tests and scales to measure mental ability. The latter measures particularly are potentially useful in proportion to their predictive value. Both types of measures have attracted an unusual amount of attention from psychologists and educators. While both are in the last analysis measures of mental processes common usage designates them as "new type educational tests" and "tests or scales of intelligence or mental ability."

## PART I

In most earlier investigations of this subject the well-known correlation technique has been employed. For the most part in earlier studies only zero order coefficients of correlation have been used. More complete analysis would point to the desirability of computing not only zero order but also partial and multiple correlations. The predictive value of the measures here considered has been ascertained by using multiple regression equations and the validity of our predictions determined by computing their standard errors of estimate. The following questions are answered in terms of our findings:

1. What relationship exists between one of our present measures of mental ability and several measures of educational achievement?

[1] The first named author is deeply indebted to Dr. R. L. Morton, Ohio University, from whose courses in statistics and helpful suggestions invaluable assistance has been derived.

2. How do objective tests of the type used herein agree with measures obtained by more subjective means, such for example as the high school and college average grades based largely on old type examinations?

3. What is the predictive value of our various measures used separately and in different combinations?

The data dealt with consist of scores and grades obtained by two hundred seven Ohio University students, chiefly freshmen. Both women and men are included in the various classes making up the total. For convenience let us designate these several measures as the following variables:

Variable 1 is the average grade made on fourteen objective tests in general psychology. These were of the three response multiple-answer type. This variable has been called $X_1$. It may interest the reader to know that psychology as studied at this university consists of three one-hour laboratory and two one-hour lecture-recitation periods a week. Woodworth's *Psychology, A Study of Mental Life* was used as a basis for lecture-recitation. The laboratory work consisted of the usual elementary experimental problems selected from Foster's manual. Statistical methods were strongly emphasized. Seven of the fourteen tests covered lecture-recitation material, the remaining seven laboratory material. The first of the test series came at the close of the third week of the semester and the remaining tests alternately lecture-discussion and laboratory each week throughout. Each test was made up of multiple-answer statements except that five of the laboratory tests contained ten such statements and in addition a statistical problem growing out of the laboratory work.

Variable 2 is the final examination score in general psychology. May's "*A Standardized Examination in General Psychology* was used. This score has been called $X_2$.

Variable 3 is the score on Form A of the *Otis Self-Administering Test of Mental Ability.* This score has been designated as $X_3$.

Variable 4 is the average grade for each student in all his college work excluding the grade in psychology for the first

semester of the college year 1925–1926. It will be known as $X_4$. The five-letter grading system was used—A B C D F. Consequently the grades were assigned numerical values. A was given a weight of five points per semester hour; B, four; C, three; D, two; and F, one.

Variable 5 is the average grade for each student's high school course. It will be known as $X_5$. Some of the grades were recorded on the student's records as letters. When such was the case A was given a value of 95; B, 90; C, 85; D, 75; F, 65.

*The zero order correlations.* The ten possible zero order correlation coefficients were calculated by the use of Ayres' modification of the Pearson formula.[2] These correlations and their

TABLE 1

*Zero order correlation coefficients (r) and coefficients of alienation (k)*

| | k | | k |
|---|---|---|---|
| $r_{12} = .491 \pm .0358$ | .87 | $r_{24} = .435 \pm .0375$ | .90 |
| $r_{13} = .410 \pm .0391$ | .91 | $r_{25} = .333 \pm .0418$ | .94 |
| $r_{14} = .584 \pm .0309$ | .81 | $r_{34} = .334 \pm .0418$ | .94 |
| $r_{15} = .280 \pm .0429$ | .96 | $r_{35} = .148 \pm .0455$ | .99 |
| $r_{23} = .338 \pm .0412$ | .94 | $r_{45} = .423 \pm .0383$ | .91 |

coefficients of alienation are given in table 1. The table is read as follows: The correlation between achievement in general psychology, as shown by the fourteen multiple-answer tests (variable 1), and the final examination grade in general psychology (variable 2) is represented by a coefficient of .491; between variable 1 and "intelligence" as measured by the Otis Self-Administering Test of Mental Ability (variable 3) by .410; etc.

Following Rugg's[3] classification we may conclude that five

[2] Ayres, Leonard P., 1920, "A Shorter Method of Computing the Coefficient of Correlations." *Journal of Educational Research*, 1: 216–221.

[3] Rugg, H. O., 1917, *Statistical Methods Applied to Education*. New York: Houghton Mifflin Company, p. 225.

of the ten zero order correlation coefficients in table 1 are indicative of a "marked" relationship. We hasten to suggest, however, that this sort of arbitrary assigning of values to correlations is somewhat questionable. Dr. Buckingham[4] has emphasized this possibility in a recent discussion. The nature of the variables involved in the investigation may determine to a much greater extent the definite interpretation to be made of a coefficient of correlation.

Another measure of correlation validity is Kelley's[5] "$k$" which tells us the per cent less accurately than pure chance our predictions would be when based on a correlation coefficient. Careful examination of the coefficients of alienation for the above correlations would indicate that even the highest ones are low in predictive value. But up to date very few zero order coefficients of correlation above .60 dealing with the relationship of educational or with mental measures have been obtained. While the zero order correlations reported by other investigators are comparable in magnitude with those reported here very few coefficients of alienation have been reported. It is of interest to note in this connection that Hull[6] has pointed out the very limited fields thus far covered in both educational and mental tests. It is highly probable therefore that we have expected too much from our tests.

The highest correlation in table 1 ($r_{14} = .584$) is interestingly significant. It is the mathematical expression of the relationship between achievement in psychology as measured by the average of fourteen objective tests and average college grade, psychology grade excluded. This finding tends to show that the objective tests are as accurate measures as is the semester average grade of the student in all his other courses. This is not as exact a statement of the validity of this test as could be

[4] Buckingham, B. R., 1927, *Statistically Thinking*. Educational Research Bulletin, 6: 230.

[5] Kelley, Truman L., 1924, *Statistical Method*. New York: The Macmillan Company, p. 173.

[6] Hull, Clark L: 1927. *The Correlation Coefficient and Its Prognostic Significance*. Journal of Educational Research, xv, 337–339.

TABLE 2

*Zero order and partial correlation coefficients*

| | | | | | |
|---|---|---|---|---|---|
| 12 | = .491 | 23 | = .338 | 35 | = .148 |
| 12.3 | = .411 | 23.1 | = .172 | 35.1 | = .038 |
| 12.4 | = .324 | 23.4 | = .227 | 35.2 | = .040 |
| 12.5 | = .439 | 23.5 | = .309 | 35.4 | = .008 |
| 12.34 | = .279 | 23.14 | = .154 | 35.12 | = − .002 |
| 12.35 | = .365 | 23.15 | = .168 | 35.14 | = .005 |
| 12.45 | = .322 | 23.45 | = .368 | 35.24 | = − .035 |
| 12.345 | = .276 | 23.145 | = .016 | 35.124 | = − .034 |
| 13 | = .410 | 24 | = .435 | 45 | = .423 |
| 13.2 | = .298 | 24.1 | = .167 | 45.1 | = .333 |
| 13.4 | = .281 | 24.3 | = .363 | 45.2 | = .328 |
| 13.5 | = .388 | 24.5 | = .334 | 45.3 | = .401 |
| 13.24 | = .225 | 24.13 | = .148 | 45.12 | = .299 |
| 13.25 | = .296 | 24.15 | = .097 | 45.13 | = .331 |
| 13.45 | = .281 | 24.35 | = .274 | 45.23 | = .328 |
| 13.245 | = .225 | 24.135 | = .078 | 45.123 | = .300 |
| 14 | = .584 | 25 | = .333 | | |
| 14.2 | = .472 | 25.1 | = .234 | | |
| 14.3 | = .520 | 25.3 | = .304 | | |
| 14.5 | = .535 | 25.4 | = .183 | | |
| 14.23 | = .436 | 25.13 | = .231 | | |
| 14.25 | = .455 | 25.14 | = .192 | | |
| 14.35 | = .476 | 25.34 | = .186 | | |
| 14.235 | = .418 | 25.134 | = .195 | | |
| 15.0 | = .280 | 34 | = .334 | | |
| 15.2 | = .142 | 34.1 | = .128 | | |
| 15.3 | = .243 | 34.2 | = .221 | | |
| 15.4 | = .045 | 34.5 | = .303 | | |
| 15.23 | = .137 | 34.12 | = .096 | | |
| 15.24 | = − .015 | 34.15 | = .123 | | |
| 15.34 | = .044 | 34.25 | = .220 | | |
| 15.234 | = − .001 | 34.125 | = − .106 | | |

*Key to variables in table 2*

1. Average of 14 objective tests in General Psychology.
2. Final Examination Score in General Psychology.
3. Score on Otis Self-Administering Test of Mental Ability.
4. Average semester grade in college excluding psychology.
5. Average High School Grade.

made if we had worked the intercorrelations or validity coefficients between the different tests. That they are as valid measures as are older and more subjective ones is quite evident from our results.

*The partial correlation.* Partial correlation coefficients permit a more complete analysis of relationship. They tell us the degree of relationship existing between two variables when we hold constant the influence of one or more other variables.

The partial correlation coefficients reported herein were worked out by Yule's[7] formula. Table 2 contains all the partials.

The correlation between variables 1 and 2 is represented by a coefficient of .491. When, however, the influence of variable 5 is held constant this relationship drops to that represented by a correlation coefficient of .439. Holding constant intelligence, variable 3, this relationship decreases to .411. If the influence of variable 4 is eliminated this same relationship is reduced to .324. Partialling out both variables 3 and 5 reduces the correlation to .365. When the influence of both variables 4 and 5 is held constant the correlation is reduced to .322 and when variables 3 and 4 are held constant the relationship is reduced to that expressed by a coefficient of .279—a figure only .003 greater than that obtained by removing the influence of variables 3, 4 and 5. And when as shown above variables 3 and 5 are partialled out, the reduction is less than by removing the influence of variable 4 alone. From the above it is apparent that high school average exerts the least influence in determining the relationship between a student's achievement on multiple-answer tests in general psychology and his final examination grade in the same subject, and that college average (variable 4) exerts the greatest influence.

The correlation between variables 1 and 3 is represented by a coefficient of .410. When the influence of variable 5 is held constant this relationship drops to that represented by a correlation coefficient of .338. Holding constant final examina-

[7] Yule, G. U., 1914, *An Introduction to the Theory of Statistics*. London: Charles Griffin and Company, 238.

tion grade in general psychology, variable 2, the correlation becomes .298, while if the influence of variable 4 is eliminated the correlation is reduced to .281. Thus variable 4 exerts the greatest single influence in determining the relationship between variables 1 and 3, while variable 5 exerts the least. When the combined influence of both variables 2 and 5 is eliminated the correlation is reduced to .296, a number higher than that obtained when variable 4 alone was held constant. Partialling out variables 4 and 5 reduces the correlation to .281, the same coefficient secured when variable 4 was held constant, showing that the influence exerted by variable 5 is so small that when it and variable 4 are held constant its reducing power is practically negligible. Holding the influence of variables 2 and 4 constant reduces the coefficient of correlation to .225, and when variables 2, 4, and 5 are held constant the coefficient is .225, exactly the same as when only variables 2 and 4 were partialled out, showing again the negligibility of the influence exerted by variable 5. The strong influence which variable 4 exerts is shown by the fact that the relationship between variable 1 and 3 is reduced to .281 when variable 4 is held constant, a figure only .056 greater than that obtained when all three variables are partialled out.

The correlation between variables 1 and 4, .584, is the highest zero order coefficient obtained. Partialling out variable 5 reduces this to a coefficient of .535, and when variable 3 is held constant it is reduced to .520. Holding constant the effect of variable 2 reduces the correlation to a number represented by the coefficient, .472, showing that variable 2 exerts the greatest influence, although this reduction is small. When both variables 3 and 5 are partialled out the correlation is reduced to .476, a coefficient higher than that obtained when variable 2 alone was held constant. Partialling out variables 2 and 5 reduces the correlation to .455, and holding constant the effect of variables 2 and 3 reduces it to .436. When all three variables, 2, 3, and 5 are partialled out the coefficient is .418. This shows that all three variables taken separately reduce the original correlation somewhat, yet it is seen that variable 2 has the greatest reducing power.

The relationship between final examination grades in general psychology and college averages is represented by a coefficient of correlation of .435. This relationship is reduced to .363 when variable 3 is held constant and to that represented by a coefficient of .334 when variable 5 is partialled out. But when variable 1 is held constant the relationship is reduced to .167 a coefficient that is only "negligible." When the combined effect of variables 1, 3, and 5 is "washed out" the correlation is reduced to .148, a number only .019 lower than that obtained when variable 1 was held constant. So it is apparent that whatever ability it is that is represented by variables 2 the same is also represented by variable 1. We should expect such to be the case as variable 2 is the final examination grade in general psychology and variable 1 is the average of the fourteen objective tests in general psychology.

The correlation between variables 1 and 4 is .584; between variables 2 and 4, .435. From the above it is apparent that the fourteen objective tests in general psychology agrees better with other college grades (variable 4) than does the final examination. Also the objective tests agree better with intelligence than does the final examination. Also the objective tests agree better with intelligence than does the final examination, for $r_{2,} = .410$. From this it would seem that little would be lost if the final examination were abolished. This conclusion does not take account of the influence of the final examination on later knowledge of and skill in psychology.

The last of our "marked" zero order correlations is that between variables 4 and 5, which is represented by a coefficient of .423. Partialling out the effect of variable 3, the correlation drops to .410. When variable 1 is held constant the relationship drops to that represented by a coefficient of .333, and when variable 2 is held constant, to a correlation of .328. Holding variables 1 and 3 constant the correlation drops to .331, a number higher than that obtained when variable 2 alone was partialled out, and when variables 2 and 3 are held constant the coefficient is .328, the same correlation obtained when variable 2 was partialled out. Holding variables 1 and 2 constant re-

duces the correlation to .299 and when variables 1, 2 and 3 are all three held constant the relation between variables 4 and 5 is .001 higher than when only variables 1 and 2 were partialled out. From this we may conclude that "intelligence" as herein measured exerts practically no influence on the relationship between college and high school averages, and that the other two variables, 1 and 2, exert very little. In other words, whatever the cause of the relationship between variables 4 and 5, it is largely foreign to our other three variables. From a prediction standpoint, when variable 4 is used as the criterion variable, this point is of significance and will be discussed later.

From the above discussion it is apparent that there are two combinations of the five variables that correlate moderately between themselves when the combined effect of the other three variables is held constant the two other combinations correlate negligibly. The moderately significant relations are those between variables 1 and 4, $r_{14.235} = .418$ and variables 4 and 5, $r_{45.123} = .300$. The other two are $r_{12.345} = .276$ and $r_{13.245} = .225$.

*The multiple correlation coefficients.*[8] Multiple correlation coefficients show the relationship between one variable called the dependent variable and the combined effect of two or more independent variables. A multiple correlation coefficient is written:

$$R_{1.234\ldots n} = \sqrt{1-(1-r_{12}^2)(1-r_{13.2}^2)(1-r_{14.23}^2)\ldots(1-r_{1n\,.\,23\ldots\overline{n-1}}^2)}$$[9]

By the use of the partial correlation technique the four highest coefficients of correlation were found to be: $r_{14.235} = .416$, $r_{45.123} = .300$, $r_{12.345} = .276$ and $r_{13.245} = .225$. These relationships were found to exist independently of the other three

[8] Yule, op. cit., p. 248.

[9] The expression, $R_{1.234\ldots n}$, refers to the multiple correlation coefficient between variable 1, the *dependent* variable, and variables 234..n, the *independent* variables. The number to the left of the point is called the *primary* subscript and those to the right of the point, *secondary* subscripts.

variables. By using the multiple correlation technique it will be shown that the independent variables rank in the same order as a basis for predicting the dependent variable as shown by the zero order correlations and in about the same order as they did when the partial correlation technique was used. For example, if variable 3 when it is held constant, causes the greatest reduction in the relationship between variables 1 and 2 it

TABLE 3

*All possible multiple correlation coefficients when variables 1 and 4 are used as the dependent variables*

| | |
|---|---|
| $R_{1.23} = .555$ | $R_{4.12} = .600$ |
| $R_{1.24} = .640$ | $R_{4.13} = .513$ |
| $R_{1.25} = .506$ | $R_{4.15} = .643$ |
| $R_{1.34} = .627$ | $R_{4.23} = .478$ |
| $R_{1.35} = .466$ | $R_{4.25} = .518$ |
| $R_{1.45} = .585$ | $R_{4.35} = .504$ |
| $R_{1.234} = .663$ | $R_{4.123} = .640$ |
| $R_{1.235} = .567$ | $R_{4.125} = .648$ |
| $R_{1.245} = .640$ | $R_{4.135} = .650$ |
| $R_{1.345} = .628$ | $R_{4.235} = .651$ |
| $R_{1.2345} = .663$ | $R_{4.1235} = .653$ |

*Key to variables in table 3*

1. Average of fourteen objectve tests in General Psychology.
2. Final Examination score in General Psychology.
3. Score on Otis Self-Administering Test of Mental Ability.
4. Average semester grade in college excluding Psychology.
5. Average High School Grade.

will add the most to that relationship when it is used in a multiple correlation.

Table 3 gives the multiple correlation coefficients with all possible combinations of independent variables when variable 1 is the dependent variable; also all possible multiple correlation coefficients in which variable 4 is the dependent variable.

The relationship between variables 1 and 2 is indicated by a zero order correlation coefficient of .491. Referring to table 3

it is seen that when the influence of variables 2 and 5 are combined the coefficient is raised to .506, and when 2 and 3 are combined the correlation rises to .555. When both variables 3 and 5 are combined with variable 2 the coefficient is increased to .567 and when variables 4 and 5 are added the coefficient rises to .640. The same result obtains when only variable 4 is added. When variables 3 and 4 are combined with variable 2 the coefficient is .663 which is identical with the correlation when all four of the independent variables are used as the secondary subscripts. From the above it is apparent that variable 5 has practically nothing to add to the relationship between variables 1 and 2. It is also evident that variable 4 makes the greatest contribution to this relationship.

In a like manner it can be seen that in the correlation between variables 1 and 3, variable 4 makes the greatest contribution when the multiple correlation technique is used and that variable 5 contributes the least. The best combination of variables in the secondary subscript when variable 1 is the dependent variable is obtained by combining variables 2, 3 and 4. Variable 5 adds nothing to this correlation. This is exactly what was found when the partial correlation technique was used.

The relationship between college average (excluding psychology), variable 4, and high school average, variable 5, as shown by the zero order correlations is .423. Turning to table 2 we see that variable 3 exerts the least influence of any single variable and variable 1 the greatest. The best combination of all the variables is that obtained by combining 1, 2, 3 and 5 which is expressed by a multiple correlation coefficient of .653.

How do objective test averages agree with traditional average grades? This question has already been answered. Partial and multiple correlations have but served to confirm the findings when zero order correlations were used. Variable 4, college average, is based for the most part on essay type tests and subjective methods of grading. Another approach to this question is to compare variables 1 and 4 with "intelligence," variable 3. The relationship between variables 1 and 3 is

represented by a zero order correlation of .410. When the influence of college and high school average is partialled out this relationship drops to that shown by a coefficient of .281. The zero order correlation between college average, variable 4, and "intelligence," variable 3, is .334 and when the influence of objective tests in psychology (14 multiple answer type tests, variable 1, and standardized objective final examination, variable 2) is held constant the coefficient drops to .96. When variables 1, 2, and high school averages, variable 5, are held constant the correlation becomes .106. If intelligence as herein measured is a criterion for judging a student's native ability it is safe to conclude on the basis of all our correlations that objective tests are the better measures. This being true, the many good features[9] of objective tests, such as saving of time in administering, added number of questions that can be given, ease of scoring, etc., are sufficient to justify their use.

May we here call the attention of the reader to yet another advantage of objective tests. Each student is enabled to plot his own progress curve as each week his test grades are given him along with the class median. Thus he has a graphic representation of how he stands in his class. This may and probably does act as a motivating factor and also may do away with the necessity of a mid-semester delinquent report.

*Prediction.* One of the greatest values of correlation is that of prediction. Such being the case what predictive value have the coefficients of correlation obtained in this study? Five variables being used theoretically there can be five criterion variables. Using in turn each of the five variables as the criterion there are 55 possible multiple correlation coefficients that could be computed. For each multiple coefficient there can be computed a regression equation. As the regression equation is the medium of prediction using each variable in turn as the criterion there is possible 55 predictions of varying

[9] Wood, B. D. *Measurement in Higher Education.* New York: World Book Company, 1923, p. 198–9; Buckingham, B. R. *Research For Teachers.* New York: Silver, Burdett and Company, 1926, p. 210–57.

values that can be made for the five variables or 11 for each variable.

In this study the regression equations were computed for the following multiple correlations only: $R_{1.2345}$, $R_{1.234}$, $R_{1.35}$, $R_{2.1345}$, $R_{3.1245}$, $R_{3.12}$ $R_{3.14}$, $R_{4.1235}$, $R_{4.35}$, $R_{4.123}$ and $R_{4.15}$.

The regression equation shows the best prediction that can be made of the criterion variable in terms of the independent variables. The standard error of estimates shows to what extent estimates made from regression equations are better than chance guesses. If we were predicting students' ability in variable 1, the best prediction we could make w thout using the regression equation would be to assign every student the known mean of that variable. In this case the standard deviation of the errors of prediction would be the same as the standard deviation of the actual scores. The standard deviation of the errors of prediction is called the "standard error of estimate." The per cent then that the standard error of estimate is of the standard deviation of the criterion variable shows to what extent the prediction is as unreliable as chance guesses. For example, consider the first case in table 4: $\sigma X_1 = 7.42$ and $\sigma 1.2345 = 5.55$ so $\frac{5.55}{7.42} = .748$ which shows our prediction of variable 1 in terms of variables 2, 3, 4 and 5 to be 74.8 per cent as bad as if it were made by assigning all students the mean score or 25.2 per cent better than pure chance.

Reading from left to right table 4 gives the multiple coefficients of correlation, their regression equations, the standard errors of estimates of our predictions based on the regression equations and the per cents better than pure chance the predictions would be.

From table 4 we find $X_1$ can be predicted 25.2 per cent more accurately than pure chance when the values of $X_2$, $X_3$, $X_4$ and $X_5$ are known. When only $X_3$ and $X_5$ are known the reliability of the prediction is 11.4 per cent better than chance. Knowing the values of $X_1$, $X_3$, $X_4$ and $X_5$, then $X_2$ can be predicted with 17.8 per cent accuracy. When only $X_1$, $X_3$ and $X_5$ are known our prediction is 17.5 per cent better than chance.

TABLE 4

*The reliability of estimates made in terms of certain regression equations*

| MULTIPLE R'S | REGRESSION EQUATIONS | STANDARD ERRORS OF ESTIMATE | PER CENT IMPROVEMENT |
|---|---|---|---|
| $R_{1.2345} = .663$ | $X_1 = .243X_2 + .144X_3 + .416X_4 - .001X_5 + 40.46$ | $\sigma 1.2345 = 5.55$ | 25.2 |
| $R_{1.234} = .663$ | $X_1 = .199X_2 + .116X_3 + .416X_4 + 43.45$ | $\sigma 1.234 = 5.55$ | 25.2 |
| $R_{1.35} = .466$ | $X_1 = .251X_3 + .322X_5 + 30.11$ | $\sigma 1.35 = 6.57$ | 11.4 |
| $R_{2.1345} = .554$ | $X_2 = .303X_1 + .011X_3 + .113X_4 + .235X_5 - 5.23$ | $\sigma 2.1345 = 6.11$ | 16.8 |
| $R_{2.135} = .550$ | $X_2 = .366X_1 + .118X_3 + .293X_5 - 6.16$ | $\sigma 2.135 = 6.13$ | 16.5 |
| $R_{3.1245} = .448$ | $X_3 = .351X_1 + .023X_2 - .228X_4 - .058X_5 + 32.95$ | $\sigma 3.1245 = 8.67$ | 10.6 |
| $R_{3.12} = .439$ | $X_3 = .421X_1 + .238X_2 + 7.99$ | $\sigma 3.12 = 8.72$ | 10.2 |
| $R_{3.14} = .436$ | $X_3 = .537X_1 + .190X_4 + 3.60$ | $\sigma 3.14 = 8.73$ | 10.0 |
| $R_{4.15} = .643$ | $X_4 = .364X_5 + .501X_1 - 35.47$ | $\sigma 4.41 = 5.65$ | 23.3 |
| $R_{4.1235} = .651$ | $X_4 = .418X_1 + .054X_2 - .049X_3 + .238X_5 - 18.68$ | $\sigma 4.1235 = 5.59$ | 24.0 |
| $R_{4.123} = .605$ | $X_4 = .457X_1 + .139X_2 + .065X_3 - 9.63$ | $\sigma 4.123 = 5.87$ | 20.4 |
| $R_{4.35} = .505$ | $X_4 = .211X_3 + .494X_5 - 21.65$ | $\sigma 4.35 = 6.36$ | 13.7 |

## CONCLUSIONS

1. The findings in this study indicate clearly that there is at best only a "marked" or moderate relationship between our present measures of educational achievement and mental ability.

2. Objective tests as used in this study agree better with college averages than do either high school averages or intelligence test scores. Our data indicate, also, that objective tests agree better with "intelligence" as herein measured than do more subjective measures.

3. In addition to agreeing better with a measure of students' "intelligence" objective tests save time and furnish a basis for progress curves which are helpful as factors of motivation. Mid-semester warnings to students are not necessary if test scores with measures of central tendency and variability are reported to students each week and the drawing of a progress curve for the class and for each class member is required.

4. If we use the final examination merely to determine the knowledge of and skill in psychology our findings would indicate that it should be eliminated. It is one of the least valuable of measures.

5. The measures of student ability as represented by the five variables in this study have no great predictive value, the best possible predictions being only 25 per cent better than chance guesses. How much better or poorer this is than prediction in the average school room or in the typical situation in the outside world we do not know.

# AN EXPERIMENT IN TEACHING PSYCHOLOGY

NORMA V. SCHEIDEMANN

*University of Southern California*

An experiment with a method for insuring effective study from students in first year psychology was conducted by the writer during the past academic year with students of the Pasadena Junior College and night classes at University College, University of Southern California. While nothing original is claimed for the method still it may be of sufficient interest to warrant its presentation. The material was worked out more carefully, presented in better form and developments watched more closely than, perhaps, would have been done if the instructor had not been making an experimental approach.

The importance of adequate daily preparations was recognized but since large class enrollments prohibited customary class discussions it was necessary to use some other method for insuring careful daily study. To meet this need as well as to keep a close check upon the work of the individual student a series of short, objective tests, twenty to thirty-five items, over each chapter of the text was devised. Sample questions of a typical test follow:

*Chapter VII. Emotions*

During anger large quantities of sugar stored in the liver are discharged by action of the (1) thyroid (2) thymus (3) adrenal (4) salivary, glands..... ——

Hunger is not an emotion because it is (1) localized (2) externally stimulated (3) unconsciuos (4) an organic state..... ——

Fatigue is classed as an (1) emotion (2) sensation (3) sentiment (4) instinct..... ——

The adrenal secretion has an inhibitory effect upon the (1) limb muscles (2) liver (3) digestive organs (4) heart..... ——

So far as is known, the organic responses are, in general, the same in anger, fear and zeal..... T F
The fact that the internal response is the same in various emotions supports the James-Lange theory..... T F
Where we find emotion we find also a tendency to action that leads to some end-result..... T F
Typically, emotion generates impulse..... T F

These tests in mimeographed form were given at the beginning of the class period. Students exchanged papers and checked the errors while the instructor read the correct responses. The giving and scoring of the tests required from five to eight minutes. Two more minutes were usually allowed for the students to look over their own checked papers and to note the errors. On the following day the distribution of class grades with the median grade was posted on a bulletin board. The following is a typical distribution.

*Chapter XI. Attention*

Third period class

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | 16 | | | | | | |
| | | | | | | | | 14 | | 16 | | | | | | |
| | | | | 9 | | | | 14 | | 16 | | | | | | |
| | | | 8 | 9 | | | 13 | 14 | 15 | 16 | | 18 | | | | |
| | | 7 | 8 | 9 | 10 | | 13 | 14 | 15 | 16 | | 18 | 19 | | | |
| | | 7 | 8 | 9 | 10 | 12 | 13 | 14 | 15 | 16 | | 18 | 19 | 20 | | 23 |
| −1 | 5 | 7 | 8 | 9 | 10 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 22 | 23 |

Highest possible score = 25
Median score = 14

Each student was thus given the opportunity of knowing his class rank at least each week. In constructing the tests[1] the important points of the chapter were selected. No attempt was made to standardize the tests in the technical sense of the term. However in the seven sections of first year psychology classes in which the tests were used there have been fairly good distributions of grades and the median grades have been quite

[1] These tests have now been copyrighted and published in collaboration with Dr. J. E. Bathurst, Bureau of Public Personnel Administration, Washington, D. C.

constant. Item difficulty was not checked. The total number of tests, twenty-one, (Woodworth's *Psychology* was used as a text) seemed large enough to permit the student to have one or two rather poor daily preparations without seriously affecting his general rank.

After several tests were given the instructor informed the students that the taking of the tests was optional with each class. Should the students prefer not to take the tests and have their class grades based upon fewer and longer tests it would meet with the instructor's approval. The classes meeting during the day were enthusiastic about continuing with the tests while two night classes asked to have the tests given, the class distributions posted, but not to have those grades recorded in the instructor's class book.

From the instructor's viewpoint these frequent short tests seemed desirable since the students seemed to study more carefully; exact meanings rather than hazy generalities were discovered. The students seemed to be better prepared to follow the development of material presented in the test as well as better prepared to follow the advanced material presented by the instructor. The elimination of subjective factors in awarding grades was also a relief. By actual check it was found that the students' rankings were quite constant. Whether the students actually did better work than if the tests had not been used was not determined since it was not possible to arrange comparable sections of students for controls.

In order to get the students' reactions to the method after they had used it for a semester a questionnaire was submitted. Whether the students were able to judge some of the points called for is neither asserted nor denied. It was thought, however, that their reactions should be considered. Typical points of the questionnaire with the tabulated responses are given below.

*Questionnaire psychology IA*

Did you study differently when you knew a short test would be given than when an assignment was made and you were certain there would not be a test?

*Studied more* 70 *Studied less* 23 *No difference* 41

Would you prefer to have the instructor include personal factors rather than be influenced entirely by the objective grades made?

*Yes* 43 *No* 85

Have you found any correlation with your grades and the amount of study you devoted to your preparation?

*Yes* 99 *No* 30

Did you study your ranking relative to the distribution of the grades made by the whole class?

*Each time* 67 *Quite often* 43 *Occasionally* 22 *Never* 2

If you studied your rank in the distribution of grades made by the whole class, was there an effect upon your study habits for subsequent preparations?

*Studied more* 103 *Studied less* 0 *No effect* 29

What has been your relative ranking in the distribution of class grades?

*About the same* 105 *Wide variations* 24

In general would you recommend the use of such tests for your work in psychology?

*Yes* 117 *No* 15

The responses to the fourteen questions presented to a total of one hundred thirty-seven students seemed to indicate that the tests were of value to that group. That student judgment is not infallible was amply demonstrated by such criticisms as the following: "the tests eliminated the factor of time in forgetting," "made one less fearful of final tests," "they put too much stress upon authority and not enough credit for the student's opinions," "grading should be absolute, not relative," and "not fair because only one-half of the students can get above the median."

# NOTES AND NEWS

## THE RELATION OF LEFT-HANDEDNESS TO PSYCHONEUROTIC TRAITS AND TO INTROVERSION

G. H. Estabrooks and E. C. Huntington

*Colgate University*

The subjects in this experiment consisted of 280 freshmen entering Colgate in the autumn of 1927. The particular measure of left-handedness used consisted in the relation between strength of grip in right and left hand. It was measured by the dynamometer manufactured at the Narragansett Machine Co., Providence, R. I. Psychoneurotic traits were measured by the Colgate Mental Hygiene Test, Personal Inventory B2 and Introversion by the Personal Inventory C2. The subjects were all male.

The dynamometer readings and certain other physical measurements are taken on all freshmen during the first week of their attendance at the university. The mental hygiene tests are also given during this period. In the case of the dynamometer the individual is first instructed as to technique and is then asked to grip the instrument once with his maximum strength. One grip only is used. Rarely do any other trials pass the first and the time factor is always important in obtaining a great variety of anthropometric data.

The degree of right-handedness or left-handedness is taken as represented by the excess in strength of grip which one hand shows over the other. This, of course, has no bearing on questions of absolute strength. Measured in this fashion, we find that the correlation—Pearson "r"—between left-handedness and magnitude of psychoneurotic traits is +.09. That between left-handedness and degree of introversion is −.146.

It would thus seem clear that this particular line of attack would yield no clear cut correlation between handedness and the traits in question.

---

A new quarterly, "Revue de la Science du Travail," has been inaugurated, the editors being Drs. Lahy, Sollier and Arend, published by the Library Felix Alcan, Paris. This publication is to be devoted to problems of psychotechnique and labor management. First, it is to be the medium for the discussion of improved diagnostic, prognostic, selection, placement and training methods in the professions as well as in

the various business and industrial kinds of work. Second, the various physiological and psychological conditions of the worker, fatigue for example, and the problems of prevention of occupational diseases and accidents will be emphasized.

The following outline indicates the scope of this new journal:

(1) Original articles relative to the scientific researches and accomplishments in the field of the science of labor; (2) general reviews of the most important questions of the day; (3) notes and documents concerning (*a*) the general progress of psychotechniques and the scientific organization of labor (methods, new procedures, statistics, forms, etc.); and (*b*) the organization and functioning of institutions of psychotechnique and the scientific organization of labor in different countries; (4) bibliography, critical analyses of books and of original articles and memoirs; (5) information relative to institutions, laboratories, journals and reviews, congresses, laws and regulations, deaths, nominations, etc.; (6) in addition there will be summaries of reviews of psychotechnique, applied psychology, physiology and hygiene of labor and the scientific organization of labor.

Dr. Emilio Mira of Barcelona, Spain, will give a seminar and lecture courses in the Department of Psychology during the coming summer session at Ohio University. The advanced course will deal with modern European psychology, particularly in its application to education. He will lecture in Abnormal Psychology and also in Vocational and Personnel Psychology. Like many other European psychologists following his summer work he will take part in the Ninth International Congress of Psychology to be held at Yale University, New Haven, Connecticut, during the first week of September, 1929.

The Stanford Department of Psychology will have as visiting professors for the summer quarter, 1929, the following: Dr. A. Michotte, of the University of Louvain, Belgium; Dr. Robert Redfield, of the University of Chicago, who will offer courses in cultural anthropology; Dr. Robert Seashore, of the University of Oregon. For the years 1927–28 and 1928–29 Mr. Richard A. C. Oliver of the University of Edinburgh has been at the Stanford Department of Psychology on a Commonwealth Fund Fellowship. He leaves at the end of this school year for Kenya Colony, East Africa, under the auspices of the British Government and the Carnegie Corporation, with authority to improve the text books in use in the native schools and to devise mental tests suitable for the selection of natives for responsible posts.

Dr. Donald A. Laird, Director of the Colgate University psychological laboratory, has been appointed technical advisor to Dr. Eric C.

Kunz, executive manager of Givaudan-Delawanna, Inc., the largest manufacturers of chemical aromas in the world with factories at Geneva, Switzerland; Lyons, France; and Passaic, N. J.

Dr. Garry C. Myers, Head of the Division of Parental Education, Cleveland College, Western Reserve University, and formerly Head of the Department of Psychology, Cleveland School of Education, has become the Editor of *Babyhood*, a magazine for parents of little children. On Friday morning, February 1, 10:20 Eastern Standard Time, Dr. Myers began a series of sixteen radio talks over WLW, Cincinnati, on home education of preschool children.

"Problem Parents" is the title of a series of thirteen articles Dr. Myers has been writing for the Child Welfare Magazine, the official organ of the National Congress of Parents and Teachers. At the request of the United States Commissioner of Education, Dr. Myers wrote a series of eleven letters to parents under the general title "Helping Our Children Succeed in School." These letters are free to all parents whose names and addresses are sent to the United States Bureau of Education, Washington, D. C.

# BOOK REVIEWS

FLOYD L. VAUGHAN. *Marketing and Advertising.* Princeton University Press, 1928. xl + 255 pp. Price $2.50.

The purpose of this book as stated by the author is to appraise marketing and advertising from the social or public standpoint, thereby rectifying the shortcomings of the typical treatment of these subjects, based upon the individualistic or acquisitive point of view. The increase in cost of marketing, the social cost and utility of advertising, acquisitive practices, coöperative marketing, and the part played by individual initiation and government action in their relation to the problems raised constitute the subject matter of the treatise. While economics rather than psychology determines the viewpoint of the book the two chapters on advertising may commend it to the attention of students of psychology of advertising.

The author dwells at some length on the change in the basis of competition from price to alleged quality or services, giving the history and economic reasons for this change. Lest any one should doubt that such a change has taken place he cites examples such as the putting up in packages and labelling of goods formerly sold in bulk. With the attempt to compete on the basis of quality it became necessary for the producer to indentify his article and to bring its superior virtues to the attention of an indifferent public. The net result has been that competition is now as keen and as disastrous in the advertising field as it ever was in the price field. But, the author asks, who pays the costs? He cites the stock arguments by which the enthusiasts for advertising attempt to show that it lowers prices and shows that in many cases these arguments apply only to individual commodities on hands. He seriously questions whether such arguments apply if the social view is adopted, i.e. whether as a result of advertising consumers as a whole pay less for what they buy. A number of specific arguments are advanced which further minimize the social gain accruing from advertising. It should be stated, however, that the author recognizes a legitimate place for advertising in the economic scheme and that his attack is based primarily upon waste in advertising and what he believes to be over-enthusiasm based upon superficial economies.

Should the author's viewpoint be accepted it is not without important implications for applied psychology. A large part of the body of doctrine of the psychology of advertising becomes a mere incident of technique for the perpetuation of an uneconomic enterprise. For, unless

we have the grace to laugh up our sleeve about what we are doing, are we not pretty thoroughly committed to the social value of advertising? While the student of "orthodox economics" will have a facile answer for many of the author's arguments, this economic appraisal of present day advertising and marketing practices must give us pause.

AMOS C. ANDERSON,
*Ohio University.*

CHARLES A. BEARD (Editor). *Whither Mankind: A Panorama of Modern Civilization.* New York: Longmans, Green and Co. 1928. Pp. vii + 408.

For many years it has been quite the fashion for our literati to lampoon "the machine age," "big business," "large scale production," and all that is linked therewith. There has existed also an obstinate tendency to applaud the values of the pre-scientific ages and to despise man's mastery of the material world. Certain foreign and domestic scoffers who write for the so-called intellectual minority insist in all sincerity (sometimes!) that Western civilization is bankrupt—that our acquisitive society is sick—and that today the nearest approach to the kingdom of heaven is to be found in the Orient.

In the volume, *Whither Mankind*, Dr. Charles A. Beard and seventeen able collaborators have attempted as dispassionately as possible to evaluate our modern Western civilization and to understand its drift. In the Introduction Dr. Beard seeks to set the problems and to give general direction to the work of his various collaborators. In style and in logic he sets a rather difficult pace. The position taken at the outset may be summarized in the following sentence:—"What is called Western or modern civilization by way of contrast with the civilization of the Orient or mediaeval times is at bottom a civilization that rests upon machinery and science as distinguished from one founded on agriculture or handicraft commerce" (p. 14). What kind of civilization is it? What are the gains and losses of the life that it offers as compared with those of other civilizations the earth has known? Science and the million machines born of science have profoundly altered the ways of men and women. What in the eyes of an intelligent realism is likely to be the future of this civilization? Is it a mere "flash in the pan," a historical accident destined to give way to some other order based upon entirely different modes of life? It is such questions as these that the book, *Whither Mankind*, seeks tentatively to answer.

Chapter I opens with Dr. Hu Shih, the Chinese philosopher, insisting that the so-called "spirituality" of the Orient is really nothing more than the religion of defeatism. Unable to cope with the forces of nature the Oriental gives up the fight in despair. Like the disappointed fox in the ancient fable who declared the grapes sour because he could not

reach them, the Oriental philosopher rationalizes his impotency by declaring that wealth and physical comforts are contemptible and that poverty and misery are something to be proud of. In almost militant style Hu Shih denies that Western civilization is less spiritual than the Eastern flight from reality. As one who knows both East and West he maintains spiritedly that the substructure of any spirituality worthy of the name is the material and economic fabric that frees mankind from the status of the savage. "What spirituality," asks Hu Shih, "is there in a civilization which tolerates such a terrible form of human slavery as the 'ricksha coolie? Do we seriously believe that there can be any spiritual life left in those poor human beasts of burden who run and toil and sweat under that peculiar bondage of slavery which knows neither the minimum wage nor any limit of working hours? Do we really believe that the life of the 'ricksha coolie is more spiritual or more moral than that of the American workman. . . . . ?" (pp. 28–29). And again he states: "Picture a civilization where boys and girls and old women with bamboo baskets tied to their backs and with pointed sticks in hand, flock to every dumping place of garbage and search every heap of refuse for a possible torn piece of rag or a half-burnt piece of coal? How can we expect a moral and spiritual civilization to grow up in such an atmosphere?" (pp. 29–30).

As Hu Shih goes about in the Far East, seeing sickness that elementary medicine could cure or prevent, starvation due to defective transportation, and appalling poverty near undeveloped resources, he cannot look with amused indifference on "well-fed persons gathered in comfortable drawing-rooms to deplore the materialism and black despair of science and the machine" (p. 405). He flatly declares that it is the machine age which rightly deserves the appellation "spiritual."

As for the condition of the present-day laborer: Although the men who today work with machinery are not always creative, joyous, or free workmen, neither was the majority of mediaeval artisans. "With respect to the mass of mankind, it is safe to assume that the level of monotony on which labor is conducted under the machine régime is by and large not lower but higher than in the handicraft, servile, or slave systems of the past" (p. 21). "When the worst has been said about the condition of the industrial proletariat, it must be conceded that as regards material welfare, knowledge, social consideration, and political power, it is far removed from the proletariat of Rome or the slaves of a more remote antiquity. . . . . a party dedicated to a deliberate return to pre-machine agriculture with its low standards of life, its diseases, and its illiteracy is beyond the imagination" (pp. 18–19).

Much of the remainder of the book is a validation of the theses stated above. Man today is on the whole and in the long run much better off than were his ancestors. The idea that there once was a golden age is a

myth. However, the authors of this volume are not oblivious to the evils of the modern order. Their optimistic attitude is always tempered. They realize that the modern slum is a fact, that abuses have crept in with the development of modern industry, that man's inhumanity to man is not lessened merely because we live in an age of steel and electricity. But they do not concede that "any other system, could it be freely chosen in place of machine civilization, would confer more dignity upon human nature, make life on the whole richer in satisfactions, widen the opportunity for exercising our noblest faculties, or give a sublimer meaning to the universe in which we labor. Quite the contrary" (p. 404).

Although the various chapters which comprise the volume under review are of very unequal merit few will fail to find among them interesting and significant observations. Since the reviewer has not sufficient skill to comment upon more than one or two of the chapters in the space allotted to him, he can only commend them to the reader with the brief statement that they are suitable for either casual or careful perusal. Indeed, it is hardly an exaggeration to say that no better effort has hitherto been made to assess and weigh the forces of twentieth century life.

HARVEY C. LEHMAN,
*Ohio University.*

PETER SANDIFORD. *Educational Psychology.* Longmans, Green and Co., New York, 1929.

Professor Sandiford has attempted to write a text from a purely objective standpoint. In the preface and the first chapter and occasionally throughout the book he reminds us that he has departed from the older and more futile methods of traditional psychology. He evidently feels the isolation of the pioneer. We welcome his book as a good addition to the ranks of educational psychologies *already* written from the objective point of view.

The author's discussion of inherited nature is largely from the biological point of view—placing more emphasis upon the Mendelian laws than is found in the average text. He reviews the usual grist of genealogies including the Jukes, the Kallikaks, and the Edwards. He also makes a considerable review of studies of twins. The point of view in the discussion is strongly hereditarian. The material on intelligence and intelligence tests is well chosen and presented. Perhaps the author gives more attention than is necessary to Spearman's theory of the nature of intelligence in view of the fact that he discusses no other except Thorndike's which he dismisses with a few lines.

The analysis of the learning process is excellent. The author makes a rather extensive study of conditioning, but exclusive of that, he fol-

lows closely Thorndike's formulation of the laws of learning. Much illustrative material is introduced from the field of animal learning. The author makes a very good analysis and criticism of the Gestalt concept of insight. A chapter on improvement in learning deals with the usual problems involved in the study of learning curves, the length and distribution of practice periods, etc. A study of the problems of retention or memory is masked under the more objective title "Permanence of Improvement." Here are reviewed experimental studies made in the field of learning and forgetting from Ebbinghaus to the present. The topic of fatigue receives a more comprehensive treatment than is found in most books. The author has very wisely refrained from reviewing all the studies (both good and poor) which have been made on the subject of transfer of training. Instead he has presented only the outstanding pieces of research and yet has furnished the student with an adequate notion of the problem.

Three chapters are devoted strictly to applied material. These deal with the measurement of instruction, and improvement in school subjects. The special subjects studied are language, reading, spelling, writing, and arithmetic. In discussing the measurement of instruction the author makes a good review of the cases for and against various types of examination questions and gives examples of improved tests. He does exhibit a faith in the accomplishment quotient which is not altogether justified by the results obtained in its use. He states that the A.Q. is a "true measure of a pupil's merit" and an "excellent measure of a teacher's efficiency." Would that it were true!

The author includes an introduction and three chapters which seem to the reviewer to belong to the field of general psychology. The introduction deals with the nature of psychology, its methods and subdivisions. The three chapters are chiefly concerned with physiology. Perhaps some portion of it may be of value to the future teacher, but the reviewer is inclined to think that more pertinent material ought to be included than the nature of sense organs and sensation.

The chapters are well summarized. Contrary to the usual custom no exercises have been included at the close of the chapters—an omission which will not disturb many readers. The author's style is clear and in spite of some technical material which is included the book is readable.

STUART M. STOKE,
*Ohio University.*

MAX SCHOEN. *The Effects of Music.* New York, Harcourt, Brace and Company, 1927. Pp. IX + 273.

The introduction by W. V. Bingham explains that this series of essays embodies the results of a competition conducted by the American Psychological Association in 1921, with a prize furnished by Mr. Edison

for the best research on the effects of music. Part I of the volume deals with different types of listeners. It includes an article by C. S. Myers reprinted from the *British Journal* in which observers related their impressions while listening passively to phonograph music. On a second presentation they wrote out their impressions in detail. Any selected portions of the record could be repeated if desirable. Different types of individuals were discovered—the associational, who for instance, visualized the music hall, etc.; the conative who tried to pick out different instruments; the sensory emotional who felt such things as a "lazy feeling," or "sting on right arm"—the same types observed in an earlier experiment with tuning forks. Unmusical observers had fewer associations. The importance of "psychic distance" or detachment is evident.

Another article in this section by Otto Ortmann approaches the matter genetically. He distinguishes three types. The first is the sensorial type, which is especially characteristic of children and musically untrained adults. Experiments indicate that medium pitch, intensity and duration are preferable. The second type is the perceptual which interprets and can be trained but with somewhat greater difficulty. The third type, the imaginative, is still more highly specialized.

Part II deals with the nature of musical enjoyment. A rather elaborate study is reported by Esther L. Gatewood. Three trained observers listened to about 600 selections played on a phonograph and filled out a rating sheet including such items as, *familiar, pleasant, unpleasant, interesting, boring, action, memory, imagination, quiet, joy*. Each item was defined on a supplementary sheet. Various correlations were computed; e.g. pleasantness and amusement correlated about .55 but pleasantness and dignity only .22. There was quite a range between the qualities showing highest relation to pleasantness and those showing lowest relation. There were indications that selections that arouse the greatest number of *different* emotional responses are the most pleasant. Other tables show the frequency with which each emotion occurs with others,—e.g. longing rarely occurs by itself while amusement often occurs alone. Individual differences between subjects are of course marked. Enjoyment is classified as physical, simple satisfaction, associational and ideational. In another study selections were used which involved especially rhythm, melody, harmony and timbre. Thirty-five subjects rated these selections on a briefer scale including however the four above-mentioned characteristics. For each selection one of the above four characteristics predominated. Moreover rhythm usually arouses happiness or excitement, while melody contributes to restfulness or seriousness. The author evidently gets out of the data about all that is to be obtained. The crux of the matter is the ability of the subjects to analyze their own emotions. They were "highly trained"

in introspection and presumably able to make such analysis. Furthermore they were required to select from a list of alternatives rather than merely describe their experiences. It seems that the experiment was as carefully designed and controlled as could be done with this particular problem and technique.

Another article in this section is by Margaret F. Washburn and George L. Dickinson. Player piano, Orchestrelle and Victrola were used to produce some 182 selections for students in a woman's college. They reported on pleasantness, source of pleasure,—rhythm, melody, harmony, etc., —exciting or quieting effect, other sources of pleasure and finally imagery. The sources of pleasure most frequently mentioned were melody, rhythm, harmony, design and tone color. This is the order of frequency of mention with melody most frequent. Composers varied in the most pronounced source of pleasure. For instance melody was the most noticeable for all but Handel, Brahms and Debussy; rhythm was first for Brahms; design ranked high for Bach and Mozart; harmony for Debussy. The same tendency is found as in Gatewood's study for pleasantness to be greater when there were a number of sources for it. This article is to be commended for its brevity and direct approach to the point.

Section III deals with the mood effects of music and includes two articles by Max Schoen and Esther Gatewood. The subjects noted on a blank their mood preceding each test and then changes in mood resulting from listening to the music. The effect of a selection with a given subject seemed to be fairly consistent. The percent of the subjects who recorded the same effect the second time varied from 45 to 91 per cent with the different selections. All but one selection had a percentage over 50. This trend is in a way a justification of the introspective technique used. Sadness, joy and rest are the feelings most frequently aroused in the experiment. Joy, amusement, sadness, stirring and rest show greatest consistency, i.e. agreement among observers, and in the above order. Emotions are more often reported with vocal than with instrumental music. The words often add another variable, but instrumental music may arouse just as intense feelings. Most individuals prefer music dynamically similar to the existing mood.

Part IV deals with the Organic Effects of Music and reprints the article of Ida M. Hyde on "Effects of Music upon Electro-cardiograms and Blood Pressure." Quantitative data are not presented in the article but it seems that music does have a definite effect on the physiological functions under investigation. There was depression in the musically susceptible for tragic minor music whereas there was general stimulation by the Toreador Song. It is suggested that music might have therapeutic value, e.g. in subduing undue hilarity or in stimulating a depressed case. In one instance lullaby on a violin inhibited an articular flutter and increased cardiovarcular tonicity in general.

Part V deals with the effect of repetition. The prize-winning essay by Margaret F. Washburn, Margaret S. Child and Theodora M. Abel is included. Sections from victrola records,—severely classical, serious popular classical, easy popular classical and popular were played five times in succession with 30 second intervals. The subjects were classified on the basis of a brief questionnaire as musical or non-musical. They rated the pleasantness of each repetition on a scale of five. There is a clear lowering of pleasantness with the popular music especially for the musical subjects. On the whole repetition increased pleasantness with the selections used except for the popular. Decreases in pleasantness seemed due to fatigue among other things. Increases in pleasantness involved imagery, comprehension of the composition, familiarity, greater attention to rhythm, melody, instrumentation or harmony, better adjustment of mood and to associations.

Another study along similar lines by A. R. Gilliland and H. T. Moore is reprinted. Classical and jazz records were played twenty-five times and measures made of tapping, strength of grip, pulse, judgment of enjoyment value and a photograph showing facial expression. The enjoyment of the classical music increased relatively with the repetition. Greater motor innervation and pulse rate was found with the jazz as well as a "becoming levity of countenance" but even from the motor standpoint repetition is more favorable to the classical music. The outstanding merit of this experiment is the effort at objective approach.

The final article in this section is by June E. Downey and George E. Knapp. The selections were classified according to national feeling, poetic thought, programme music and formal construction. They were given at intervals of a week for five weeks and the order systematically varied in each program. Repetition tended to increase pleasantness within the limits of the experiment (no popular or jazz music was included). The poetic thought group was the most preferred in general and the more subtle compositions profited most from the repetition. Some selections had more imagery than others. Familiarity tended to increase along with the pleasantness and sometimes the imagery was influenced by knowledge of the nature of the selection, e.g. that it was Chinese.

The last section discusses effects other than auditory and organic. There are further articles by Ortmann and Gatewood. We note such things as the desire to see the concert pianist's hands and the use of similes in describing music. With some selections there is considerable uniformity in the use of such similes.

The book is not, of course, a systematic and unified treatment of the subject, involving as it does a series of somewhat miscellaneous essays. The editor does inject comments at the beginning of each section—possibly slightly too enthusiastic comments. There is also naturally some

duplication as the studies were made independently, but this is no serious objection and has a certain merit in that the studies corroborate one another on certain points. There is much statistical material included and evaluated perhaps as critically as the type of data warrants. There is no mention, for instance, of probable error in the book although differences between averages are sometimes stressed.

As a whole the work is rather heavy for the general reader but of course will not be so for one with a research interest in the field. It would have helped greatly if the editor had made an effort at the end to write a rather critical summary of the whole series of essays. This might have been a rather herculean task because there is such a wide range of material covered. The book shows the difficulty of the experimental approach to the effects of music and it is evident that we are dealing with something that is not particularly "getatable." Objective techniques can be used at present for only a limited portion of the problems. We must rely on highly trained observers or else amass a tremendous amount of data for untrained observers and use a rigorous statistical analysis.

If on another occasion anyone desires to subsidize research on the effects of music, the reviewer feels that instead of offering a "prize for the best research" it would be better to organize a systematic program and provide fellowships for each project in that program.

HAROLD E. BURTT,
*Ohio State University.*

## NEW BOOKS AND PAMPHLETS RECEIVED*

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

*Annotated Bibliography Dealing with Examinations and School Marks.* C. W. O'DELL, University of Illinois, Urbana, Ill. 50 cents. 42 pp.

*Enrollment in the Foreign Languages in Secondary Schools and Colleges of the United States.* CARLETON A. WHEELER AND OTHERS. Macmillan Company, New York. 453 pp.

*Free Association Test.* GRACE H. KENT AND AARON J. ROSANOFF. John Wiley & Sons, New York City. $1.50. 620 pp.

*Freud's Tragischer Komplex.* CHARLES E. MAYLAN. Verlag Ernst Reinhardt in Munchen. 215 pp.

*The Hygiene of Instruction.* LAWRENCE AUGUSTUS AVERILL. Houghton Mifflin Company, New York City, 1928, 386 pp.

*Methoden der Experimentellen Psychologie.* Abt. VI. Urban & Schwarzenberg, Freiderichstrasse 105, Berlin, N24, Germany. 192 pp.

*Modern Psychology.* GARDNER MURPHY. Harcourt Brace and Company, New York. $4.50. 470 pp.

*Race Attitudes in Children.* BRUNO LASKER. Henry Holt and Company, New York. 1929. $4.00. 394 pp.

*Revue de Psychologie Concrete.* DIRECTEUR C. POLITZER. "Les Revues" 47 Rue Monsieur-le-Prince, Paris. 157 pp.

*Social Psychology.* CARL MURCHISON. Clark University Press, Worcester, Mass. $3.50. 210 pp.

*Summary of Investigations Relating to Grammar, Language, and Composition.* R. L. LYMAN. University of Chicago, Chicago, Ill. 302 pp.

*University of Iowa Studies in Psychology—Seashore Commemorative Number.* Edited by WALTER R. MILES AND DANIEL STARCH. Psychological Review Company, Princeton, N. J. 223 pp.

*Wahrheit Und Wirklichkeit.* DR. OTTO RANK. Franz Deuticke, Wien I, Helferstorferstrasse 4, Leipzig, Germany. 112 pp.

*Youth and The Good.* School Betterment Studies, Vol. I, No. 2. HENRY C. FRICK. Educational Commission, 469 Union Trust Building, Pittsburgh, Pa. 76 pp.

---

* Mention here does not preclude further comment.

## Directory of American Psychological Periodicals (*Continued*)

**Journal of Personnel Research**—Baltimore: Williams & Wilkins Co. Subscription $5.00. 500 pages approximately. Edited by Walter V. Bingham, and Clarence S. Yoakem. Bi-Monthly.

**Genetic Psychology Monographs**—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison and an international co-operating board. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

**Psychological Abstracts**—Princeton, N. J. Subscription $6.00. 600 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

**Mental Measurement Monographs**—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Founded 1925. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research.

**Studies in Psychology and Psychiatry**—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Founded 1926. Edited by Edward A. Pace. Published without fixed dates, each number a single research.

**Journal of General Psychology**—Worcester, Mass.: Clark University Press. Subscription $7. 600–700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

## Directory of American Psychological Periodicals

**American Journal of Psychology**—Ithaca, N. Y.; Morrill Hall.
Subscription $6.50. 600 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

**The Pedagogical Seminary and Journal of Genetic Psychology**—Worcester, Mass.: Clark University Press. Subscription $7. 700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

**Psychological Review**—Princeton, N. J.: Psychological Review Company.
Subscription $5.50. 480 pages annually. Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

**Psychological Bulletin**—Princeton, N. J.: Psychological Review Company.
Subscription $6.00. 720 pages annually. Psychological literature.
Monthly. Founded 1904. Edited by Samuel W. Fernberger.

**Psychological Monographs**—Princeton, N. J.: Psychological Review Company.
Subscription $6.00 per vol. 500 pp. Founded 1895. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches.

**Psychological Index**—Princeton, N. J.: Psychological Review Company.
Subscription $4.00. 300-400 pp. Founded 1895. Edited by W. S. Hunter. An annual bibliography of psychological literature.

**Journal of Philosophy**—New York; 515 W. 116th St.
Subscription $4. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge and Wendell T. Bush.

**Archives of Psychology**—Sub-Station 84, N. Y.: Archives of Psychology.
Subscription $6. 500 pp. ann. Founded 1906. Ed. by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.

**Journal of Abnormal Psychology and Social Psychology**—Albany, N. Y.
Subscription $5. Boyd Printing and Publishing Co. Edited by Morton Prince in co-operation with Floyd H. Allport.
Quarterly. 432 pages annually. Founded 1906. Abnormal and Social.

**Psychological Clinic**—Philadelphia: Psychological Clinic Press.
Subscription $2.50. 288 pages. Ed. by Lightner Witmer. Founded 1907.
Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

**Training School Bulletin**—Vineland, N. J.: The Training School.
Subscription $1. 160 pp. ann. Ed. by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.

**Comparative Psychology Monographs**—Baltimore: Williams & Wilkins Co.
Subscription $5. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.

**Psychoanalytic Review**—Washington, D. C.: 3617 10th St., N. W.
Subscription $6. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

**Journal of Experimental Psychology**—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental.
Subscription $6.00. Founded 1916. Bi-monthly. Ed. by Madison Bentley.

**Journal of Applied Psychology**—Baltimore: The Williams & Wilkins Co.
Subscription $5.00. 400 pages annually. Founded 1917. Bi-monthly.
Edited by James P. Porter, Ohio University, Athens, Ohio.

**Journal of Comparative Psychology**—Baltimore: Williams & Wilkins Co.
Subscription $5. 500 pages annually. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

*(Continued on Third Cover)*